Covert Contrast in the Acquisition of English /ɹ/: A Case Study

by
© Josh Roberts

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

September 2019

St. John's, Newfoundland and Labrador, Canada

Abstract

In this thesis I examine the potential role of covert contrast (the phenomenon by which a phonological contrast produced by a language learner falls below the threshold of perception of adult speakers of this language) in the development of one child learning the contrast between /ɹ/ and /w/ in English. More specifically, using a longitudinal corpus documenting the development of American English by one child learner, I compare data obtained by impressionistic means of phonetic transcription against acoustic measurements of the same speech tokens. As we will see, the results from both the transcription data and the acoustic measurements mirror one another in ways that undermine the claim that covert contrast represents a necessary stage in acquisition (Scobbie et al. 1996). Additionally, the current study reveals a disparity in the time of acquisition for /ɹ/ in coronal stop-initial onsets vs. all other onset environments which highlights the influence of articulatory factors on the production of /ɹ/ across different phonological contexts. Finally, the acoustic component of the study uncovers what can be termed a 'covert allophone' in the case of /w/ in that same coronal-initial complex onset environment. I conclude this thesis with a discussion of both the theoretical and methodological implications for future research on the development of phonological contrasts by children.

Acknowledgements

To Dr. Yvan Rose, for guiding me, assisting me, putting up with me, keeping me on track, reading and re-reading this thesis, supporting me even when I couldn’t support myself, and being the best advisor I could have asked for, je vous remercie mille fois. I would not have been able to do this without you and I am eternally grateful for your guidance.

To Drs. Sara McKenzie, Paul De Decker, Sandra Clarke, and Bob Hollett, the experience I have earned working under you is irreplaceable. Thank you for your leadership.

To the entire linguistic faculty at Memorial University, you have been a second home for years now, and as I journey away from Newfoundland I am confident that you have given me the preparation necessary to continue in this field. Thank you.

To my fellow students at Memorial University, thank you for the moral and social support. Without you, I would have probably been a shut-in for most of my time here. Particular thanks to Naomi and Alysha, whose work dovetailed with and contributed to my own.

To my parents, Tamar and Randy Roberts, my brother Jared, and all of my family, you have supported me in every way imaginable over the course of this thesis and my life, and I’m sure you will continue to be there for me in the future. Thank you, and I love you.

To my partners, Callum Sizer and David Ferris, you have been my rocks throughout this entire process. Thank you, I love you, and I look forward to many, many more years with you.

To Irma Gerd, Eda Kumquat, Liezel Hues, /garbagefile, Wych Hazel, and Madame Daddy, this past year has been the best of my life thanks to you all. I'll miss you. Keep the garbage goblin spirit alive; I know I will be.

To Andrea Szeszko, ✨🖤✨

Table of Contents

List of Tables

List of Figures

Chapter 1: Introduction

The language of children has long been a topic of study in linguistics, primarily for the fact that the speech of children patterns differently from that of adults in many ways, some of which make it at times not easily comparable to adult speech (Jakobson 1941; Ferguson 1964; Smith 1973; Priestly 1977; Rose & Inkelas 2011; McAllister Byun, Inkelas & Rose 2016; Rose 2017). Recent methodological developments in the field, as well as in technology more generally, have allowed for more finely-tuned analysis of how and why children behave the way they do during the acquisition of their language. One of the benefits of these developments, modern acoustic analysis, has allowed us to observe developmental behaviours that would have gone unnoticed when using traditional methods, for example based on (impressionistic) phonetic transcription. The addition of an extra, non-human set of 'ears' in deciphering acoustic data has also raised questions about human-only approaches to data descriptions. Addressing this methodological debate, I examine one common pattern of child speech, commonly referred to as '/ɹ/-gliding' or 'rhotic gliding', to determine how, and to what extent, what could be perceived as a mispronunciation on the child's part may instead relate to limitations on data interpretation based upon human, impressionistic transcriptions of child speech.

1 Child speech

The speech of children typically displays many patterns of error, some of which we do not observe in adult speech. Consider a child who makes the following speech errors (data from Ingram 1974):

(1) Velar fronting (Ingram 1974:240)

a.	/keɪk/	→	[teik]	“cake”
b.	/ki/	→	[ti]	“key”
c.	/gɹin/	→	[diːn]	“green”

Although the productions in (1) are incorrect when compared to their target forms, they are not random. In each case, the child is taking a velar stop and articulating it as its alveolar counterpart. (In (1c), the stop is part of a cluster which is simplified to a singleton onset; this simplification is independent from the pattern described.) The examples in (1) illustrate a common pattern, known as velar fronting. Certain child error patterns are not as straightforward as the one in (1), in which every velar onset is fronted to its alveolar equivalent. The data in (2a) below, obtained from a different child, suggest that this child regularly pronounces /l/ as [j], whereas (2b) shows that he can also pronounce /l/ as [w].

(2) Gliding of /l/ by E, an L1 learner of English (Inkelas & Rose 2007:712–714)

a.	/vijəˈlɪn/	→	[vijəˈjɪn]	“violin”
	/lɪdɪə/	→	[ˈjɪdɪə]	“Lydia”
b.	/hɨldə/	→	[hɨwdə]	“Hilda”
	/æləgetɚ/	→	[æwədɛɾɚ]	“alligator”

Although certain analyses may take this variable pattern as evidence that the child is behaving randomly with regards to the articulations of target consonants (e.g. Hale & Reiss 1998), the phonological environments for each misarticulation are, in fact, different: the articulation of /l/ as [j] occurs in prosodically strong environments (onsets of word-initial or stressed syllables), whereas [w] productions occur in prosodically weak ones (syllable codas, onsets of unstressed syllables) (Inkelas & Rose 2007:713).

Types of systematic child speech errors fall into two groups overall. Consider the difference between the errors in (3a) vs. (3b):

(3) Child speech errors with and without adult analogue (Rose & Inkelas 2011:4–6)

a. Child speech errors with adult analogue

i.	Syllable reduplication:	/ɛləfənt/	→	[fæfæ]	“elephant”
ii.	Vowel epenthesis:	/bluː/	→	[bəluː]	“blue”

b. Child speech errors without adult analogue

i.	Consonant harmony:	/teɪbɫ/	→	[beːbu]	“table”
ii.	Consonant fusion:	/smoʊk/	→	[fok]	“smoke”

While the patterns in (3a) have correlates in adult languages, those in (3b) are unique to child speech. Syllable reduplication, for example, is attested in many languages such as Tagalog, Kham, and Zande (Blake 1917; Watters 2009; Pasch 2017). Vowel epenthesis occurs in many languages as well, including in situations where words with consonant clusters are borrowed into languages that do not allow clusters, such as the English word “strike” adapted into Japanese as “sutoraiku” (Rose & Demuth 2006; Shoji & Shoji 2014). While the errors in (3b) are not attested in adult speech,[1] they are typically systematic (Rose 2000:17–18): consonant harmony in child languages targets the major places of articulation of a consonant and changes it to that of another consonant which it is in proximity to, and consonant fusion takes phonological properties from two adjacent consonants and combines them into a single sound. The selection of which phonetic features are triggers or targets of this process is dependent on a child’s individual phonological system. As such, these examples cannot be analyzed as instances of lexical substitution; they reflect systematic outputs of the child’s developing phonological system.

1 Certain languages do display consonant harmony, but consonant harmony in adult languages never involves major place of articulation; generally, it involves only minor featural changes affecting [±anterior], [±lateral], etc. (Young & Morgan 1987; Tryon 1995; Hansson 2010).

However, certain productions patterns involve variation which is harder to characterize phonologically, as they also involve clear conditioning at the level of speech phonetics and articulation. A canonical case of this is commonly referred to as “covert contrast”, which I discuss further in the next section.

2 Covert contrast

Covert contrasts, a term coined by Hewlett (1988:31), are “impressionistic homophones which are acoustically or through articulation different” (Scobbie et al. 1996:44).[2] This acoustic or articulatory difference implies that when a speaker produces a contrast between two sounds, this contrast may not be perceived as such by the listener. For example, if articulations of /ɹ/ which are perceived as [w] are different from articulations of target /w/ also perceived as [w], these articulations involve a covert contrast, since the contrast which exists between the two sounds (target /ɹ/ vs. target /w/) is not perceived by the average listener. Covert contrasts have been found in many types of non-standard speech, such as child speech, disordered speech, and the speech of second language learners (Churchill 2009; Lin & Demuth 2011).

As suggested by Scobbie et al. (1996), covert contrasts in child speech could arise when children are perceptually aware of a contrast between two phonemes but lack the motor control required to reproduce the relevant distinctions in their speech productions. As Macken & Barton (1980) and Scobbie et al. (1996) suggest, there may indeed be a lag between the time a child begins to learn how to pronounce a sound and the time when the articulation of this sound is fully mastered by this child. During this period, the child’s attempted articulations of a given phone may be perceived as corresponding to a different phoneme in the language, as in the example above, in

2 Other terms have been used for this phenomenon, including “pseudo-neutralization” and “subphonemic contrast” (Stokes & Ciocca 1999).

which /ɹ/ is perceived as [w]. From a theoretical standpoint, the presence or absence of certain types of covert contrasts may shed light onto the mechanisms through which children acquire phonological contrasts in production, and the degree to which child speech errors may be due to articulatory problems or to the representation of phonological contrasts in the child's mental lexicon, among other possible explanations. On a practical level, the potential existence of covert contrasts suggests that a child's actual phonological abilities may be initially underestimated for phonetic, articulatory reasons (i.e. motor issues affecting speech production). Through the ability to perform an analysis that verifies or eliminates covert contrasts as a reason for the child's errors in production, we are able to obtain clearer pictures of a child's actual degree of phonological development for different speech sounds.

3 Thesis overview

In this thesis, I examine one behaviour, referred to as /ɹ/-gliding (or rhotic gliding). /ɹ/-gliding is a pattern whereby a child produces a target consonant /ɹ/ in a manner such that it is perceived as a [w] by adult listeners. This pattern is robustly attested in the literature on L1 English learners (McGowan, Nittrouer & Manning 2004; Richtsmeier 2010; Rose & Inkelas 2011). It has been suggested (for example by Richtsmeier 2010) that this behaviour is in fact not a substitution of /w/ for /ɹ/, but rather a case of covert contrast. My findings, using acoustic analysis of the formant structure of target /ɹ/, glided /ɹ/, and target /w/, show that there is no noticeable difference between the acoustic patterns of a glided /ɹ/ and a "true" /w/. However, I also find an interesting behaviour in the development of /ɹ/ in coronal-stop-initial complex onsets: the child I studied masters /ɹ/ in this environment one year before he is able to produce it correctly in any other environment. The child also does not produce a /tw/ onset as [tɹ], which provides evidence for there being some sort of representational difference in his mind between /ɹ/ and /w/, even if this difference is not perceived by adult listeners or detected using acoustic analysis in all other contexts.

In Chapter 2 of this thesis, I provide an overview of the previous literature that my research builds upon, with a focus on research dealing with covert contrast involving the /ɹ/~/w/ distinction. In Chapter 3, I discuss my research project, the questions I examine through my research, and the methodology of my study, including data selection. Chapter 4 describes the development over time of /ɹ/ and /w/ in this child's speech, through the lens of adult listeners' impressionistic transcriptions of his speech. I examine the same tokens in Chapter 5, where, rather than using transcriptions by adult listeners, I use acoustic measurements. These two developmental analyses and their implications are discussed in Chapter 6 along with implications of my methodology and findings for the field at large; I conclude with some final thoughts. Appendix A and Appendix B contain all of the individual graphs which comprise the more comprehensive graphs I describe in Chapters 4 and 5, and the standard deviations for my formant measurements, respectively.

Chapter 2: Background Literature

In this chapter, I provide a brief overview of the literature on covert contrast. I examine the literature on both /ɹ/-gliding and other situations in which covert contrast has been found to occur in both adult and child speech. There is a wealth of literature on covert contrast, in particular in the context of child speech development (Macken & Barton 1980; Young & Gilbert 1988; Tyler, Edwards & Saxman 1990; Scobbie et al. 1996; Carter & Gerken 2004; Churchill 2009; Richtsmeier 2010; McAllister Byun, Buchwald & Mizoguchi 2016). I begin with a chronological survey of the literature pertaining to covert contrast, starting with early analyses based on legacy methods, and moving into modern examinations of covert contrast across various speech contexts. Following this, I continue with a summary discussion on why important questions remain despite the wide reach of past examinations of the topic, and how these questions can be addressed empirically.

1 Early analyses of covert contrast

Kornfeld & Goehl (1974) offer, to my knowledge, the earliest examination of potential covert contrast in child speech. Although Kornfeld & Goehl did not rely on instrumental analysis, their examination nevertheless suggests the presence of covert contrasts affecting child speech. Several children, all of whom glided /ɹ/ to [w], were recorded by Kornfeld and Goehl while producing /ɹ/~/w/ minimal pairs (e.g. “red”~“wed”, “right”~”white”). These recordings were then played back to both adult listeners and the children themselves, who were asked to choose which samples involved the words beginning with /ɹ/ or /w/. While the adults could not correctly distinguish the children's productions of /ɹ/ from those of /w/, the children were far more reliable in selecting the correct words from the recordings: they identified /ɹ/ words from their own recordings at a better-than-chance rate, and selected the correct /w/ words with near-perfect accuracy. This ability displayed by the children, but not the adults, suggests the existence of acoustic differences between the two sounds produced by the children that the adults were unable to perceive. This research was thus a precursor to the idea of covert contrast, and offered a basis for research on the phenomenon based on child speech development, up to and including the present study.

Macken & Barton (1980), in their study of child speech phonetics, examined the stages that children exhibit in their development of voicing contrast among obstruent stops. The researchers recorded four children's natural speech, starting at around age 1;06, at 16 points over the course of eight months. Macken and Barton then isolated tokens of pairs of voiced and voiceless consonants (such as /p/~/b/, /k/~/g/) from these recordings. Through spectrographic and oscilloscopic analysis of these recordings, as well as impressionistic descriptions, Macken & Barton uncovered three separate stages in development. The first stage was characterized by the complete absence of any voicing, with neither an impressionistically perceptible contrast in the production of voiced vs. voiceless consonants, nor a contrast measurable through instrumental analysis. Skipping over to the third (and final) stage, adult listeners could perceive a contrast in the children's productions, which was also reflected in both the spectrograms and oscillograms. However, in the intermediate stage, there was a disconnect between the adult listeners' perception and the data derived from instrumental analysis: while adult listeners were unable to identify the children's productions of voicing-differentiated target phones as being different sounds, acoustic measurements of the same tokens revealed a difference between the children's /p/ and /b/, /k/ and /g/, etc., for the majority of the pairs of sounds produced by all four children. This disconnect between adult perception and acoustic analysis provides evidence for the presence of a covert contrast: a voicing contrast was measurably present during this intermediate stage which adult listeners were unable to perceive.

2 New hypothesis

Casting a view towards covert contrast as a stage in acquisition, Scobbie et al. (1996) examined the stages of development between two sets of sounds in the speech of a developmentally delayed child. This child, code-named DB, was 4;1 at the outset of the study and was undergoing speech therapy at the time. DB initially produced two sets of sounds, /t/~/d/ and /st/~/d/, with no perceived difference, but matured into producing a measurable (but covert) contrast in later sessions. The researchers performed instrumental analysis of the VOT of DB's productions of these consonants across the developmental period. In the later recording sessions where there was no still perceptible contrast between /t/ and /d/, there was a measurable difference in breathiness, which the researchers took to be a proxy for voicing.

While Macken & Barton (1980) described three developmental stages, Scobbie et al. (1996) suggested four stages in the development of a phonological contrast, two of which are evidenced by their data.

(4) Scobbie et al.'s (1996) four stages of contrast development

a. No contrast
b. Covert contrast
c. Immature contrast
d. Mature contrast

In the first stage, the child does not produce any noticeable contrast between the two sounds either in perception or through instrumental analysis. In the second stage, there is a covert contrast in the child's productions of these sounds, as shown in Scobbie et al.'s study by the systematic and measurable difference in the breathiness of the productions of /t/ vs. those of /d/. (/st/ patterned with /t/ in DB's speech.) Despite this measurable difference, the consonants were not perceived by listeners as belonging to different phonemes. The cues for this contrast in DB's speech were particularly elusive, as the difference between /t/ and /d/ was not marked by VOT, which distinguishes the sounds in adult speech. The lack of adult-like cues during this period in DB's speech makes this contrast an "inappropriate contrast" (Scobbie et al. 1996:44). Such contrasts are difficult to both perceive and study because they are not detectable using the expected phonetic correlates of the adult contrast. The last two stages that Scobbie et al. (1996) discuss, but which are not exemplified in their data, are a stage of "immature contrast", in which the contrast is perceptible by listeners but not realized in an adult-like (or target) manner, and finally the mastery stage, in which the child displays full command of the adult forms, producing a "mature contrast" (Scobbie et al. 1996:44).

One implication under a strong interpretation of Scobbie et al.'s (1996) proposal is that *all* children undergo the four developmental stages listed above in (4), at least to some extent. However, given the high level of variability observed even in the speech of typically-developing children, this strong claim may be untenable. It is indeed possible that at least some children do not display covert contrasts at all, or skip some of the stages proposed by Scobbie et al. (1996); thus, covert contrast may not be an obligatory stage in phonological development. The research I introduce in the next chapter indeed suggests this weaker interpretation of Scobbie et al.'s (1996) proposal.

3 Additional evidence for covert contrasts

3.1 Deletion

Carter & Gerken (2004) investigated the phenomenon of unstressed syllable deletion in child speech productions. Previous research has shown that, in words with unstressed initial syllables, children frequently delete the unstressed initial syllable and begin the word with the stressed syllable, for example reducing /bəˈnænə/ to [ˈnænə] (Carter & Gerken 2004:562). Carter & Gerken analyzed child productions with deleted initial syllables and compared them to target forms prosodically similar to the children's output forms. They elicited sentence pairs from children who reliably deleted unstressed initial syllables. These pairs of sentences, although different in their target forms, were perceived as near-identical by adult listeners due to the syllable deletion (e.g. "He kissed ~~Lu~~cinda" versus "He kissed Cindy"; "He pushed ~~Cas~~sandra" versus "He pushed Sandy") (Carter & Gerken 2004:567). Acoustic analysis of these sentences showed that when the children were deleting an initial syllable, the onset consonant (always an /s/ in these examples, e.g. "~~Lu~~[s]inda") was lengthened in comparison to the forms in which no syllable deletion occurred (such as "[s]indy"). This lengthening suggests that the children had some mental representation for the truncated syllable; despite the full form not being realized in their output, the shape of the target word form influenced their speech patterning in an acoustically measurable fashion.

Acoustic traces of deleted elements are also observable from simplified consonant clusters in the speech productions of Dutch-learning children (Gulian & Levelt 2009; Gulian & Levelt 2011; Gulian 2017). Gulian & Levelt (2009) examined cases of deletion in /tr/- and /kn/- initial clusters, both of which are common complex onsets in Dutch and frequently reduced by children to a single consonant. Gulian & Levelt analyzed an existing corpus of longitudinal Dutch child speech data, as well as data recorded from three additional children between the ages of 1;11 and 2;09, who also displayed the cluster reduction patterns. These children were recorded while repeating a list of words which either contained the cluster expected to be reduced or the comparable singleton onset in a similar phonological environment (e.g. “klippen” /klɪpən/ vs. “kippen” /kɪpən/, “trein” [trɛɪn] vs. “tijd” /tɛɪt/). For the /tr/ vs. /t/ onsets, Gulian & Levelt performed acoustic analysis on the change in formant values from the consonant to the vowel. Gulian & Levelt analyzed the /kn/ vs. /k/ onsets for the nasal murmur, a set of formants and one antiformant characteristic of nasality in a spectrogram. They found a significant difference in the height of F3 between the reduced /tr/ clusters and the singleton /t/ onsets, but not between the formant patterns of the reduced /kn/ clusters and the singleton /k/ onsets. However, while the difference was not measurable within their chosen parameters for /kn/ vs. /k/, Gulian & Levelt did note a visible difference in the spectrograms between the two types of onsets, which they suggest is a difference of vowel quality (Gulian & Levelt 2009:8–9). They hypothesized from this the existence of a covert contrast, implying that the full extent of the children’s phonological knowledge is obfuscated by the inability of adult listeners to perceive it in all of its detail.

In a follow-up study, Gulian & Levelt (2011) analyzed cases of simplification affecting word-initial /sC/ clusters. They recorded 12 Dutch children between 1;08 and 2;08 years of age producing words that begin with /sC/, /s/, and /C/ onsets, which were paired with each other in groups based upon the quality of the following vowel. For example, the test word “staart” [sta:rt] correlated to the control words “taart” [ta:rt] and “saap” [sa:p]. The children produced the words in an elicitation task in the carrier phrase “twee _______”, where the blank indicates the plural form of the word after the children were presented with the singular. The researchers separated the test tokens into groups between those in which the first consonant was deleted (C1) vs. those in which the second was deleted (C2), and performed acoustic analysis on all tokens. For the C1 tokens, they measured the time between the preceding vowel (i.e. that of “twee”) and the onset of the initial consonant. For the C2 contexts, they measured the time between the onset [s] and the vowel of the word. Comparisons between the test and control tokens revealed a statistically significant difference in the time between the preceding vowel and consonant for the C1 tokens, and between the [s] and the following vowel for the C2 tokens. These findings reveal leftover traces of the deleted elements for what impressionistically appears to be neutralization through deletion, and suggest, in line with Carter & Gerken (2004), that the children were retaining some representation of the adult target form, as reflected in acoustic assessments of their productions. Taking a careful stance on these results, Gulian & Levelt leave room for further research regarding whether the reduced form or the full adult form is represented in the children’s lexicon. (Also see Gulian (2017) for further discussion of cluster reduction acquisition in Dutch.)

3.2 Velar fronting

Regarding the difference between a target /t/ and a fronted /k/ (to [t]), research has been inconclusive with regards to the existence of a covert contrast. In a study by Tyler, Edwards & Saxman (1990), half of the four children studied displayed a VOT distinction between a target /t/ and a fronted /k/, whereas no contrast was found in Young & Gilbert’s (1988) study. In a study on stop bursts (energy created with the release of a stop), Forrest et al. (1990) found that three children studied did not display a systematic difference between the two consonants, while a fourth child did display a significant difference between target /t/ and target /k/. Building upon this inconclusive body of work, McAllister Byun, Buchwald & Mizoguchi (2016) set out to examine a potential covert contrast using a different method: ultrasound imaging. In their examination of two children who regularly fronted the velar /k/, as well as of two children who produced the mature contrast between /k/ and /t/, McAllister Byun et al. measured the VOT of target /t/ and /k/ produced by the child participants and took ultrasound images of the children's lingual articulation of these sounds. For all children, the VOT measurements were not significantly different between target /t/ and fronted /k/ in the children who fronted velars. The ultrasound measurements were inconclusive as well, with one child not displaying a contrast in their tongue positioning and the other displaying one. However, the child who did not display a contrast fronted velars more frequently, and the child who did display a covert contrast matured into producing an overt contrast during later recording sessions. These findings may suggest that the first child was at the “no contrast” stage of development, and that the second child at the covert contrast stage, following the stages suggested by Scobbie et al. (1996) and laid out in (4) above. However, the results may also suggest that covert contrast is not a universal phenomenon in development, as only one child was found to display it. These findings are significant for my research, for two reasons. First, they add to the question as to whether covert contrast is a necessary stage in phonological development, in line with the general hypothesis by Macken & Barton (1980) and by Scobbie et al. (1996), since the child who did not display a covert contrast did not appear to have moved directly to an adult-like contrast, at least based on the available data. Secondly, these findings show that a covert contrast may only be noticeable through certain means of measurement, as ultrasound detected a contrast in one child’s productions, whereas acoustic analysis did not.

3.3 Gliding

Churchill (2009) performed an acoustic case study of the four approximants (/ɹ/, /w/, /j/, and /l/) in the speech of an adolescent, code-named Marshall, who had been diagnosed with childhood apraxia of speech. (For the purposes of this literature review, I focus on Marshall’s articulations of /ɹ/ and /w/ only.) Consider, for example, the following minimal or near-minimal pairs:

(5) /ɹ/~/w/ neutralization, impressionistically transcribed (Churchill 2009:71–72)

a. Minimal pair

i.	/wɛn/	→	[wɛn]	“when”
ii.	/ɹɛn/	→	[wɛn]	“ren”

b. Near-minimal pair

i.	/wɛt/	→	[wɛt]	“wet”
ii.	/ɹɛd/	→	[wɛd]	“red”

Each of the pairs in (5a) and (5b) presents a contrast between onset /w/ and /ɹ/ in the adult (target) forms. However, in Marshall’s speech, each word in the pair was perceived by adult listeners to begin with a [w]. Churchill performed acoustic analysis on the fundamental frequency and formant values of tokens taken from Marshall’s speech for minimal pairs such as those listed in (5). She found that the acoustic traces of these approximants were actually significantly different. For every formant value, there was a significant difference in the values for target /w/ vs. /ɹ/. Additionally, there was a significant difference between the formant values (F3-minus-F1, F3-minus-F2, and F2-minus-F1), which can serve as a metric to describe formant patterns (Churchill 2009:52). These differences both in individual formant values and general formant patterning for both target sounds imply that Marshall was producing these sounds in different ways, even if the differences between them could not be discerned by adult listeners.

In a study of typically-developing children, Richtsmeier (2010) examined the production of several sounds and phonological contrasts for which children reliably make phonemic errors. Richtsmeier argues that these patterns of production do not simply consist of substitutions of one sound for another, but rather that these errors involve covert contrasts, something that is unaccounted for in many analyses which rely solely on impressionistic transcription. One of the patterns of error that Richtsmeier examines is /ɹ/-gliding, a contrast which is “extremely difficult” to transcribe and which may be hiding under “subtle acoustic differences” (Smit 1993). In order to analyze these sounds without relying solely on potentially erroneous transcriptions, Richtsmeier utilized ultrasound data obtained from two children who consistently glided target /ɹ/. He found that while the two sounds were impressionistically identical to adult listeners, the children used reliably different tongue shapes in their productions of each of the two sounds, and were thus producing a covert contrast between them.

Klein et al. (2012) addressed differences in the perception of /ɹ/ using two groups of adult listeners, one of which was experienced with regards to children’s speech and one of which was less experienced. Using data drawn from a previous study of child speech (Klein, Davidson & Grigos 2009), two certified speech-language pathologists (deemed the ‘experienced listeners’) first ranked the tokens of /ɹ/ using a three-point scale, with 1 indicating an adult-like production and 3 indicating a different phoneme. The researchers then presented the same data to a group of 12 graduate students in linguistics (the ‘inexperienced listeners’), and asked them to perform the same ranking scale on the data after three hours of training. Following this ranking, Klein et al. performed acoustic analysis to ascertain the formant values for the tokens. Their analysis shows that the ratings given by the experienced speech-language pathologists did correspond to a measurable difference in /ɹ/ or /w/ productions, and that the speech-language pathologists were accurate overall in their judgment. However, the ratings given by the inexperienced listeners did not match up with those ratings given by the experienced listeners.[3] Additionally, the ratings given by the graduate students varied in how much they matched up with the ratings given by the speech-language pathologists, ranging from 66% agreement for one graduate student to less than 20% for another. One of the factors affecting the students’ agreement with the speech-language

3 The results from the current study are very much in line with these findings; see further discussion in Chapter 6.

pathologists for a particular token was how target-like the token was: a token which was rated as adult-like by the experienced listeners was 81% likely to be rated the same by the inexperienced listeners, whereas one that was rated a 3 on the scale (i.e. perceived as a different phoneme) by the experienced listeners was only 61% likely to be rated the same by the inexperienced listeners. This disconnect between the sound implied by the token's formant values and the sound that the inexperienced listeners believed it was suggests that impressionistic interpretations of child speech (particularly with regards to the /ɹ/ vs. /w/ contrast) may not be representative of the articulated sound, in particular if this interpretation is not supplied by an experienced listener. I examine this relation with my own data (although my method is not the same as the one used by Klein et al. 2012), through a systematic comparison of a set of impressionistic transcriptions against acoustic values measured from the same tokens. As we will see, careful phonetic transcription performed by experienced transcribers do appear to find correlates in acoustic measurements.

4 Summary

Covert contrast is a well-studied phenomenon involving a phonological distinction produced by a non-standard speaker which can be uncovered instrumentally but which does not reach the threshold of human phonological discrimination. Researchers such as Macken & Barton (1980) and Scobbie et al. (1996) suggest that covert contrast is a universal phenomenon in language acquisition, a hypothesis that I test with my research, through the longitudinal case study described in Chapter 3. Other studies have shown that covert contrasts manifest themselves in many cases of phonological neutralization, including the neutralization of distinctions between singleton and complex onsets, and in the context of weak syllable deletion. Covert contrasts have been found using a variety of analytic methods: for example, while Richtsmeier (2010) examines a covert contrast in /ɹ/~/w/ neutralization in children using ultrasound images, Churchill (2009) studies the same contrast in the speech of a teenager with apraxia of speech through spectrographic formant analysis.

Although the studies summarized above do create a strong body of research (which I intend to build upon and contribute to), a *longitudinal* case study regarding the emergence of the /ɹ/~/w/ contrast has yet to be performed. Such a study is fundamental to testing the idea that covert contrast is a stage in acquisition, as the existence of a period of covert contrast must be examined based on the change in one (or many) speaker's productions over time, something that cross-sectional studies cannot provide (Rose & Inkelas 2011).

Chapter 3: Methodology

Through the research presented in this and subsequent chapters, I examine the development of the /ɹ/~/w/ phonological contrast in English, based on a longitudinal case study. In particular, I focus on the following research questions:

(6) Research questions:
   a. Is covert contrast in /ɹ/-gliding a necessary stage in child speech development?
   b. How does the /ɹ/~/w/ contrast evolve over time, both qualitatively and quantitatively?
   c. What are the implications of this study for research on child phonological and phonetic development?

First, per (6a), I examine whether covert contrast is always present in speech development by providing evidence from a longitudinal case study. In order to address (6b), I use the longitudinal data to examine how the development of the /ɹ/~/w/ contrast manifests itself over time. Finally, as stated in (6c), the results of my study serve as a basis for discussion of the implications for research on child phonetic and phonological development. In order to facilitate this study, I use a mixture of human interpretations (through transcriptions) and instrumental analyses (of formant values) of a corpus of child speech which displays regular patterns of /ɹ/-gliding, as shown by extant phonetic transcriptions of this child's speech productions.

1 Data selection

I am drawing upon data from the English-Providence corpus, which has been the topic of both phonological and morphological analyses of child language in past research (e.g. Demuth, Culbertson & Alter 2006; Song, Sundara & Demuth 2009; Evans & Demuth 2012). This corpus documents the linguistic development of six children in the Eastern United States from the ages of 1 to 4, recorded in spontaneous, naturalistic interactions with their caregivers. The English-

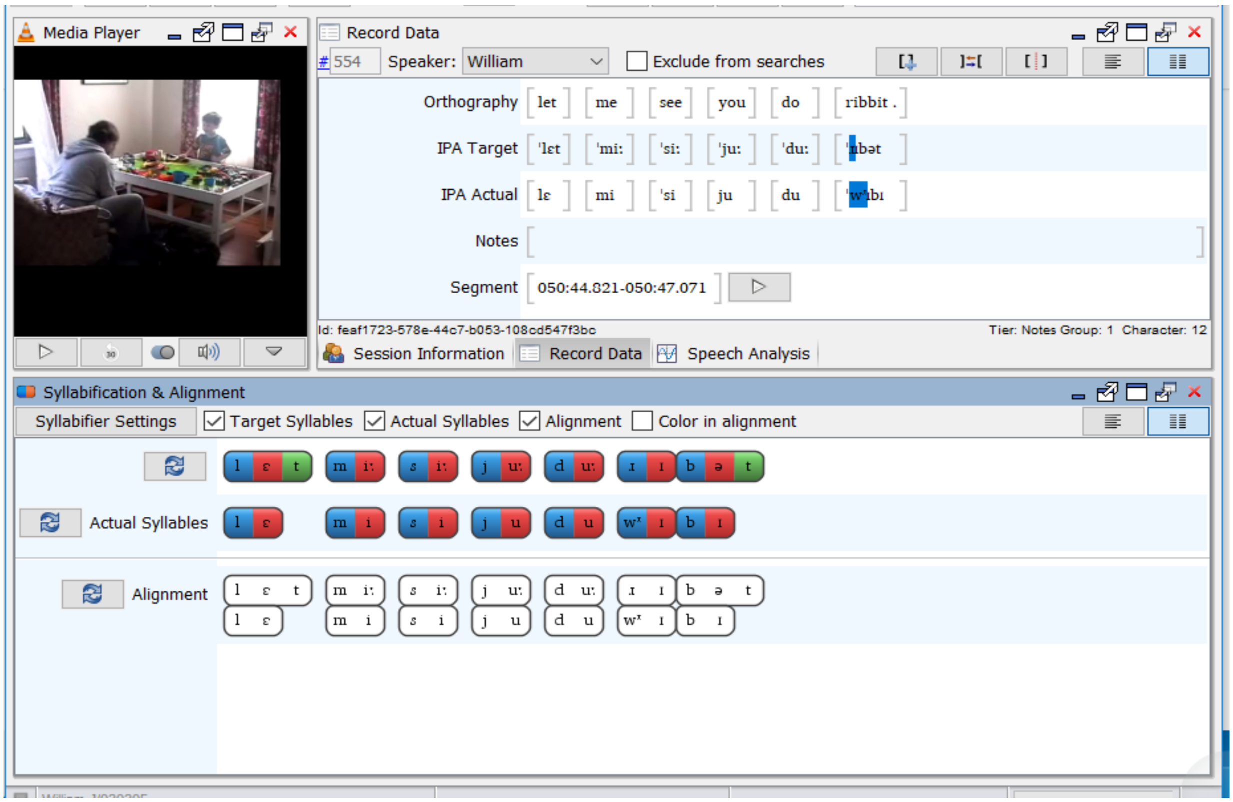
Providence corpus is available through PhonBank (https://phonbank.talkbank.org; Rose & MacWhinney 2014), an online database of corpora documenting phonological development across languages and speaker populations. My study uses longitudinal data from one child in the Providence corpus, named William, who robustly displays the apparent behaviour required for the proposed research: during the initial period covered by the corpus, William's productions of /ɹ/ are variable; during a second stage, his onset /ɹ/s are perceptually glided to [w]; and finally in later sessions, his productions of /ɹ/ are transcribed as adult-like. To the best of my knowledge, there also has not been published research on rhotic gliding or covert contrast based on this particular dataset.

2 Data preparation

Using Phon, a database software program designed to facilitate tasks related to transcription data and acoustic analysis of speech (Rose et al. 2006), I sorted through William's dataset to find records in which William produced words containing either onset /ɹ/ or onset /w/ (or both), and performed a phone alignment check. The purpose of the alignment check is to verify that the target and actual phones are paired correctly in the corpus for the analysis of segmental production patterns. Working with another graduate student in linguistics, we performed a systematic verification of the extant transcriptions of all occurrences of target /ɹ/ and /w/ found in syllable onsets in the corpus, as well as a correction of other errors noted in the transcription of the words that contained these sounds. During this verification step, we performed further annotations of the relevant consonants: transcribing a sound as [ɹ] indicates the transcribers' perception of an adult-like rhotic sound; while [w] represents a sound perceived as fully glided, [w$^{\text{ɹ}}$]/[ɹ$^{\text{w}}$] stand in for sounds between the two target phonemes, with the main consonant indicating which of the two the sound produced most resembled (Munson et al. 2010).

Furthermore, during this step a number of tokens (n=1703) were excluded, mostly due to their being unsuitable for acoustic analysis, as they were affected by audio clipping or hiccuping, or interference by another speaker or noise.[4] An example of the records edited in the process is in Figure 1.

Figure 1: Phon record during the editing of the transcriptions and alignment in the corpus



This doubly-verified version of the corpus serves as a basis for my analysis of William's phonological development based upon impressionistic transcriptions presented below, and provides a starting point for my acoustic analysis of /ɹ/ and /w/ tokens. To analyze William's

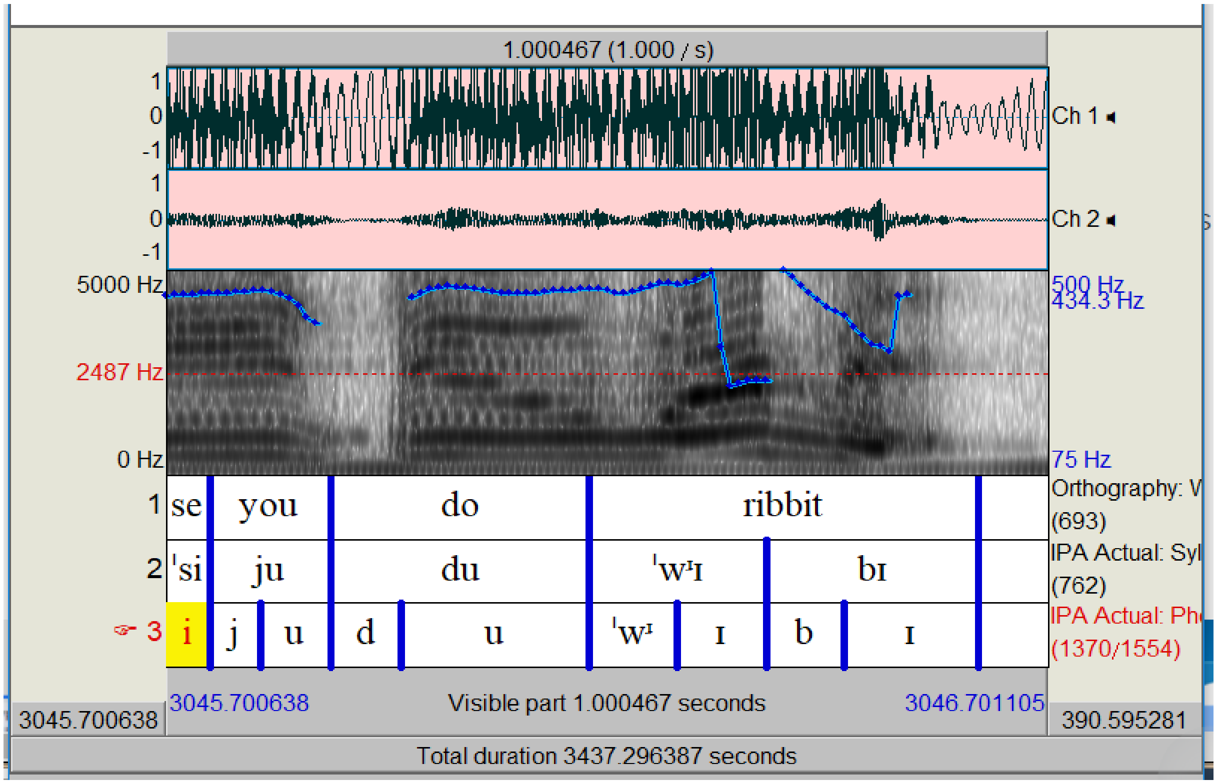
4 Although these tokens were excluded from both the developmental and acoustic analyses, this does not change the overall results as the overall developmental pattern remains essentially the same and does not impose undue influence on the consonant of covert contrast. In fact, this work on the corpus will only help us address the questions at hand, both using IPA transcriptions and based on acoustic measurements. The point that there is no covert contrast between a glided /ɹ/ and a target /w/ is immaterial to this data clean-up.

development of /ɹ/ and /w/ based on the transcribed data, I generated bar graphs detailing the number of tokens which were produced as target, intermediate, glided, substituted, or deleted for onset /ɹ/; and target, substituted, and deleted for onset /w/. (For the purposes of these developmental charts, "substituted" means that the sound was produced as a phone which was not perceived on the [ɹ]~[w] spectrum for /ɹ/, and a sound which was not perceived as [w] for /w/.[5]) These charts are described and analyzed in Chapter 4.

I analyzed these same tokens using the Praat functions for acoustic analysis (Boersma & Weenick 2017) built into Phon. In preparation for this analysis, I generated TextGrids for the tokens which contain an onset target /ɹ/ or /w/ produced by the child and aligned the TextGrids to the relevant phones in the tokens. The following is a depiction of the TextGrid alignment of phones to their spectrographic renditions.

5 The transcribed corpus contains a single instance of /w/→[ɹ] and no instances of other rhoticizations of /w/; I interpreted this as a single aberration as opposed to any systematic pattern of behaviour.

Figure 2: TextGrid being generated and aligned in Praat (Boersma & Weenick 2017)



After generation and alignment of the TextGrids, I performed acoustic analysis on the /ɹ/ and /w/ tokens, obtaining the F1, F2, and F3 values of these particular sounds. (Although normalization is necessary in many studies of vowel formants in order to eliminate variation which may arise due to physiological differences between speakers, in a one-speaker study such as the current one, normalization is not required; Disner 1980, Chambers 2008.) To obtain these measurements, I took the formant values for each token at 50%, 60%, and 70% of the sound's duration and averaged them. If individual values varied greatly within one formant for a single token (for example, an F2 of 1300 Hz at 50%, 2800 at 60%, and 1400 at 70%), the single outlying value was excluded from this average. I then generated an overall average of each formant (F1, F2, F3) for each sound (rhotic /ɹ/, glided /ɹ/, /ɹ/ produced as an intermediate sound, and target /w/) in each relevant phonological environment (singleton onsets; complex onsets beginning with labial, coronal, or velar sounds). I then performed a comparison between the /ɹ/ tokens which were judged adult-like, the target /w/ tokens, and the /ɹ/ tokens which were deemed glided, to determine if the glided tokens of target /ɹ/ pattern with the target /w/ tokens or have different acoustic traces, in order to determine the existence of a covert contrast in William's speech. (Overall, there were not enough intermediate tokens in any one environment to draw any conclusion therefrom.) I plotted these average formant values into line graphs using LibreOffice Calc. In order to smooth the graphs and obtain a more easily interpreted picture of the overall behaviour of William's development more accurately, the formant values were grouped from the 44 original sessions into 22 time points, each time point containing the values for tokens taken from two consecutive sessions, recorded across an approximately 2-week time interval.[6] This comparison and the resulting graphs are discussed further in Chapter 5.

3 Comparison baselines: /ɹ/ in adults and children

In order to measure both the values and the similarity to adult forms of the target /ɹ/ and /w/ tokens, it is beneficial to have baseline values against which to compare. To that end, I summarize in this section a selection of papers regarding the acoustic values and general spectrographic patterns of both adult and child /ɹ/ and /w/ in English.

3.1 Adult /ɹ/

Prevocalic /ɹ/ in adult English speech is characterized acoustically by the low height of its third formant (F3) when compared to that of the surrounding vowels, with a mean value of 1600 Hz, while the F3 of surrounding vowels or other sounds may be between 2100 and 3000 Hz (Espy-Wilson 1992). In addition, the rhoticity of /ɹ/ is marked by the relative closeness of F2 and F3

6 Note that individual monthly periods (1;05, 2;10, and 3;02) are missing from the data. This is simply an artifact of the temporal spacing between some of the original recordings.

(and thus a lower F3-minus-F2 value), which can help distinguish these sounds from similar glides and semivowels, including /w/. However, /ɹ/ in other positions (syllabic, postvocalic) has different acoustic properties from an onset /ɹ/ (Boyce & Espy-Wilson 1997). This is evident when comparing the formant values from each position for adult productions of /ɹ/ and /w/, listed in Table 1 and Table 2 below.

Table 1: Mean formant values for adult /ɹ/ in multiple positions (adapted from Espy-Wilson 1992)

	F1	**F2**	**F3**
Prevocalic /ɹ/	419	1285	1779
Intervocalic /ɹ/	460	1240	1720
Postvocalic /ɹ/	503	1300	1830

Table 2: Mean formant values for adult /w/ prevocalically and intervocalically (adapted from Espy-Wilson 1992)

	F1	**F2**	**F3**
Prevocalic /w/	381	848	2320
Intervocalic /w/	349	771	2340

Although a majority of the research done on the acoustic traces of these sounds has been done on adult speakers (e.g. Espy-Wilson 1992; Boyce & Espy-Wilson 1997; Espy-Wilson et al. 2000), a few researchers have focused on the acoustic patterns of English /ɹ/ in children, which I address in the next section.

3.2 Child /ɹ/

Dalston (1975) examined the acoustic traces of proper articulations of /ɹ/, /l/, and /w/ in child speech. Although Dalston only looked at articulations judged "correct" by adult listeners, the acoustic values presented by Dalston (1975) still provide a foundation for many of the empirical comparisons I will perform as part of my analyses. Tokens of word-initial /ɹ/ and /w/ (as well as /l/) were elicited from 10 Midwestern children, age 3;03 to 5;04, with an approximate mean of 4;0. Dalston then performed acoustic analysis on these tokens to obtain their formant values. The mean values of the children's /ɹ/ and /w/ are summarized in Table 3 below.

Table 3: Mean F1, F2, and F3 values for children's productions of /ɹ/ and /w/ (adapted from Dalston 1975)

	F1	**F2**	**F3**
/ɹ/	431	1503	2491
/w/	402	1020	3547

The pattern of formant values is in line with the pattern observed in adult speakers, a result possibly influenced by the fact that these measurements were based on tokens whose articulations were *a priori* judged as correct by human listeners.

McGowan, Nittrouer & Manning (2004) report solely on the acoustic production of /ɹ/ in children who are acquiring the pronunciation of that sound. They considered longitudinal data from nine Midwestern children who, starting around age 1;0, were recorded at two-month intervals until the children's mean length of utterance was over 3 words and contained regular use of function words. From these sessions, McGowan et al. extracted tokens of /ɹ/ in all possible positions (i.e. syllable initial and final /ɹ/, as well as syllabic /ɹ̩/ in both medial and final positions), based on words that "would have contained an [ɹ] or a syllabic [ɹ] if spoken by an adult with a rhotic dialect" (McGowan, Nittrouer & Manning 2004:6). These utterances contained deviant articulations of /ɹ/ towards the beginning of the observation period, and more adult-like articulations as the children's pronunciation improved. McGowan et al. analyzed the moment of the articulation most "/ɹ/-like", using spectral analysis, as well as the midpoint of the neighbouring vowel for F2 and F3 values. They found that for all syllable placements of /ɹ/, F2 and F3 decreased with age, but that prevocalic /ɹ/ had both higher formants and a larger F3-minus-F2 value than other syllable placements of /ɹ/. The value is approximately 2000 Hz, even up to the end of the sessions at 31 months, with an F3 of 3700 Hz and an F2 between 1600 and 1900 Hz at 26-31 months. These /ɹ/ measurements from the final sessions are summarized in Table 4.

Table 4: Mean F2, F3, and F3-minus-F2 for prevocalic, postvocalic, and syllabic /ɹ/, taken from last sessions measured (adapted from McGowan, Nittrouer & Manning 2004)

	F2	**F3**	**F3-minus-F2**
Prevocalic /ɹ/	1600-1900 Hz	~3700 Hz	~2000 Hz
Postvocalic /ɹ/	1900-2400 Hz	~3200 Hz	~1000 Hz
Medial /ɹ̩/	~2200 Hz	~3200 Hz	~1000 Hz
Final /ɹ̩/	2100-2600 Hz	3000-3600 Hz	800-1300 Hz

Considering that adult-like /ɹ/ is marked by the closeness of F2 and F3 to distinguish it from similar sounds such as /l/ and /w/, these data suggest that to learn the target articulation of prevocalic /ɹ/ is more difficult for children, who persist in deviant articulations of /ɹ/ in this position longer than in other syllable positions. (A verification of this possibility transcends the scope of this thesis.)

Klein et al. (2013), similar to their 2012 study summarized in Chapter 2, addressed the differences between typically-developing children and children with phonological impairment in the production of /ɹ/. Their study provides a basic framework for examining typically-developing children's erroneous productions of /ɹ/. Two speech-language pathologists initially rated the productions of /ɹ/ for both typically- and atypically-developing children on a scale of 1 to 3, the same scale used by Klein et al. (2012), discussed in Chapter 2, Section 3.3 above. The researchers then took formant measurements from these tokens at the lowest F3 point, and compared those which were judged to be adult-like with those that were deemed different phonemes. These “different” tokens displayed higher F3 values as well as a larger difference between F3 and F2. In addition, the children with phonological impairment, who were undergoing speech therapy for their /ɹ/ articulations, decreased the distance between F3 and F2 over the course of the observation period. Klein et al. concluded from this that the F3-minus-F2 value represents a robust acoustic feature for /ɹ/, with the F3 values considered a side effect from the smaller distance between the two formants.

The formant values and patterns provided by the above studies offer baselines against which to compare the articulations of /ɹ/ in the data introduced below, for example, to determine whether the formant values obtained point to adult-like /ɹ/, /w/-like productions of this consonant, or neither. More generally, these values also serve as validation for those obtained through the current study, something essential given that child speech imposes challenges to acoustic analysis, particularly because of the distance between harmonics in high-pitched voices and the high level of variability in both pitch and voice common to child speech (Buder 1996).

Chapter 4: Analysis based on impressionistic transcriptions

In this chapter, I detail the apparent development of /ɹ/ and /w/ in William's speech in both singleton and complex onsets, as analyzed through phonetically (impressionistically) transcribed data by trained linguists. These data are presented in graphs, organized by type of onset (singleton and complex) and the place of articulation of the initial consonant for complex onsets. For more detailed breakdowns of the data, organized by position within the word, stress, and initial consonant voicing in the case of complex onsets, please consult Appendix A.

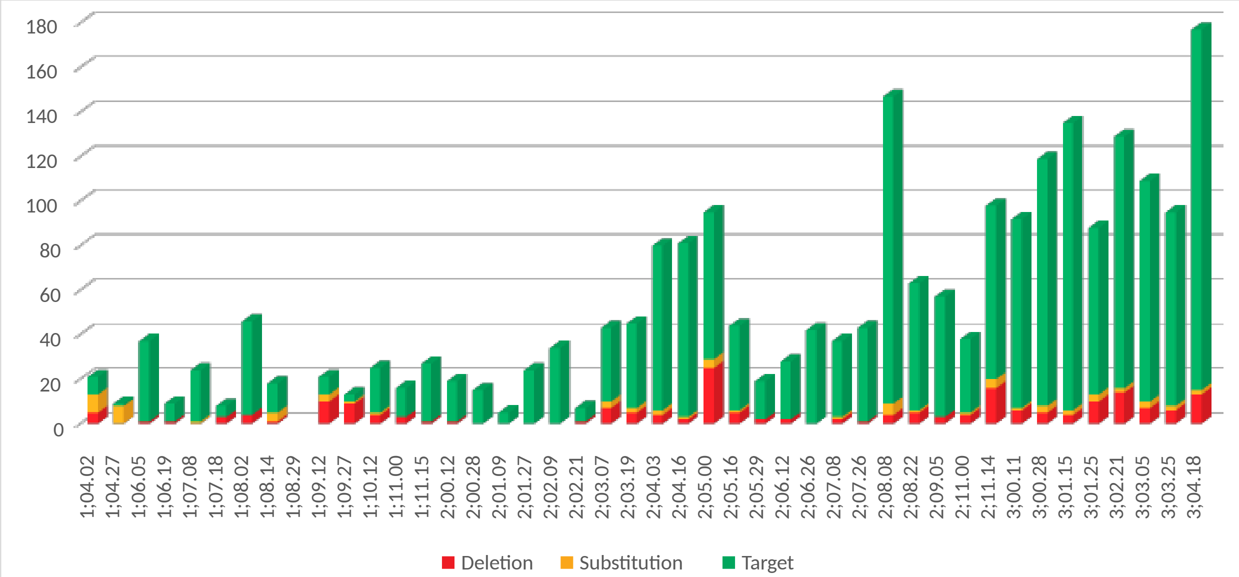
1 Singleton onsets

I begin with a description of William's productions of /ɹ/ and /w/ in singleton onsets, starting with the latter which, through all descriptions, serves as a baseline against which to compare the former.

1.1 /w/ in singleton onsets

Consider the data in the following graph. As we can see, William's articulation of /w/ in singleton onsets appears generally target-like from the beginning of the observation period, a pattern which represents the majority of his productions already as of 1;06.05.

Figure 3: Development of singleton /w/ onsets



Thereafter, only two sessions (1;09.12 and 1;09.27) deviate from this norm. The deletion patterns in these two sessions almost entirely comes from two individual words, “welcome” and “wanna”. William reduces these lexical exceptions in early sessions to the monosyllabic “come” and a single vowel respectively. Two later sessions, 2;05.00 and 2;11;14, also contain a large number of deleted instances of onset /w/. These deletions are also almost entirely in the context of the word “wanna” (or occasionally “want”); however, William articulates this word correctly as [wanə] (or as [wan]) in a majority of tokens. This variability suggests that the deletions which occur here are either due to misarticulations, or influenced by additional factors such as prosodic effects. (This question is immaterial to the current discussion.) More generally, the small number of substitutions present (particularly in the earlier sessions) are not ascribable to any pattern. The early acquisition of this sound stands in contrast to the articulations of /ɹ/ in singleton onsets, described below, which underwent periods of high variability and apparent gliding before reaching an adult-like pronunciation.

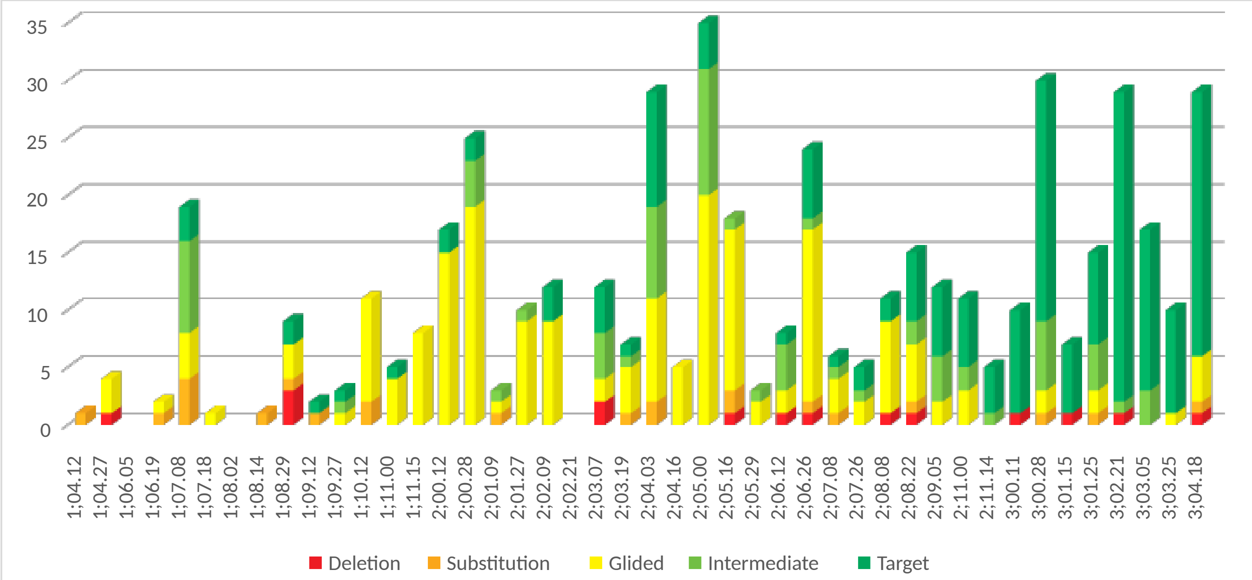
1.2 /ɹ/ in singleton onsets

As we will see in the following graphs, articulations of /ɹ/ in singleton onsets in William's speech are regularly glided at approximately 1;11 years of age, depending on the position of onset /ɹ/ within the word. From there, the articulation of singleton /ɹ/ in William's speech becomes adult-like in a majority of cases at approximately the same age across all environments (around 2;09). Articulation remains consistently adult-like from this point, with the exception of individual sessions, in which target articulations are still the most common type of articulation spoken by William. Sections 1.2.1 and 1.2.2 detail the patterning of singleton onset /ɹ/ in word-initial and word-medial positions, respectively.

1.2.1 Word-initial singleton /ɹ/

Figure 4 illustrates the development of word-initial singleton /ɹ/ in William's speech, as impressionistically parsed by adult transcribers.

Figure 4: Longitudinal development of word-initial singleton /ɹ/

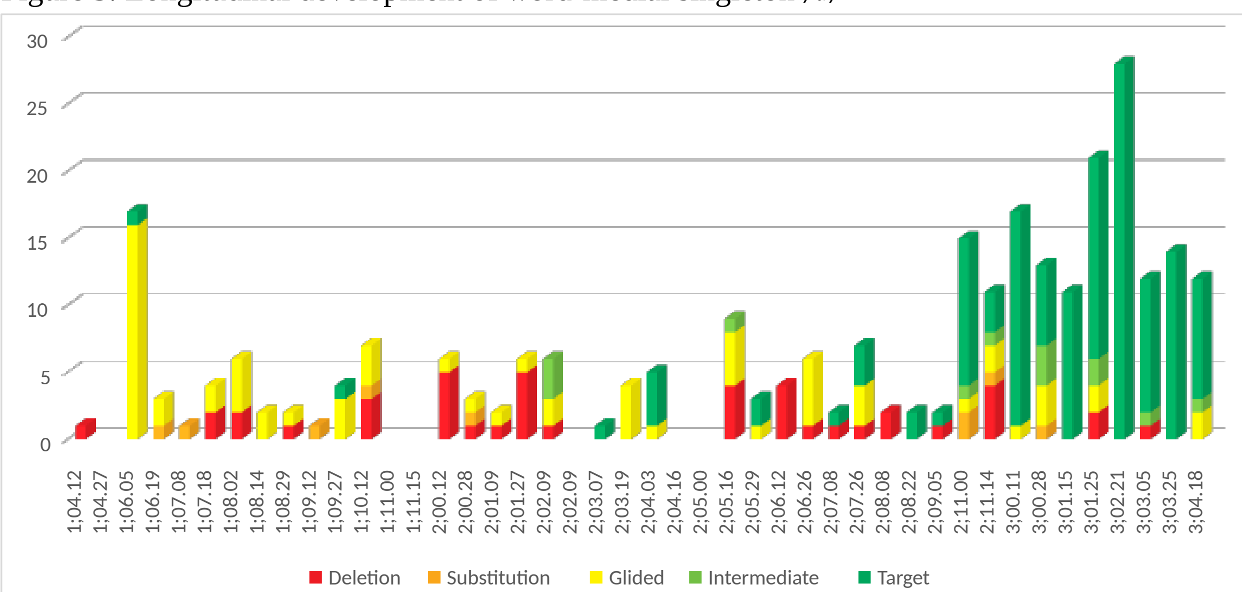


In the earliest sessions (those prior to 1;10.12), William's articulations of /ɹ/ are highly variable. We observe, for example, a high incidence of substitutions at 1;07.08, and deletions at 1;08.29. Starting at 1;10.12, glided articulations become the vast majority of apparent articulations, until age 2;03.07. From then on, glided articulations remain a prominent pattern, especially when considered alongside the intermediate articulations, until the point in time at which William's target-like articulations become the majority. This point of mastery occurs at around age 2;11.00; target-like articulations also occur in greater-than-chance amounts in certain earlier sessions (e.g. 2;04.03, 2;06.26), and in large numbers in the few sessions immediately before 2;11.00.

1.2.2 Word-medial singleton /ɹ/

There are considerably fewer tokens for word-medial /ɹ/ onsets in the data. Despite the lower number of tokens, we can still clearly observe patterns throughout the data, illustrated below in Figure 5.

Figure 5: Longitudinal development of word-medial singleton /ɹ/



William's acquisition of target-like pronunciations occurs largely at the same time as for word-initial /ɹ/ (i.e. around 2;11.00). The main difference between word-initial and word-medial articulations prior to mastery is the high occurrence of deletions of word-medial /ɹ/ during sessions where initial /ɹ/ was almost invariably glided or semi-glided. However, these deletions occur primarily in productions of the words "orange" and "camera"; outside of these two words, gliding comprises the most prominent articulation of word-medial singleton /ɹ/ in William's speech until target-like articulation of /ɹ/ is achieved.

2 Complex onsets

Similar to the above, this section details William's development of /ɹ/ and /w/ in complex onsets, considering also the major place of articulation of the first consonant in the target cluster. Due to a difference in patterning between /ɹ/ in complex onsets which begin with a stop consonant and in those which begin with a fricative (particularly with coronal-initial onsets), these environments are discussed separately. Sections 2.3 through 2.5 describe /ɹ/ in onsets which begin with a stop, and section 2.6 describes /ɹ/ in onsets which begin with a fricative. This division was not necessary for /w/ in complex onsets, as /w/ patterned the same no matter the manner of articulation of the initial consonant in the onset.

For all complex onsets containing an /ɹ/, the majority of productions before age 1;09 result in /ɹ/ deletion. This pattern is likely due to onset cluster reduction driven by William's overall phonological pattern (e.g. Smith 1973; Fikkert 1994; Goad & Rose 2004) rather than any phonetic behaviour specific to /ɹ/, and thus I do not explore this patterning in depth in this description.[7] Keeping with the organization of my data description in Section 1, I begin with

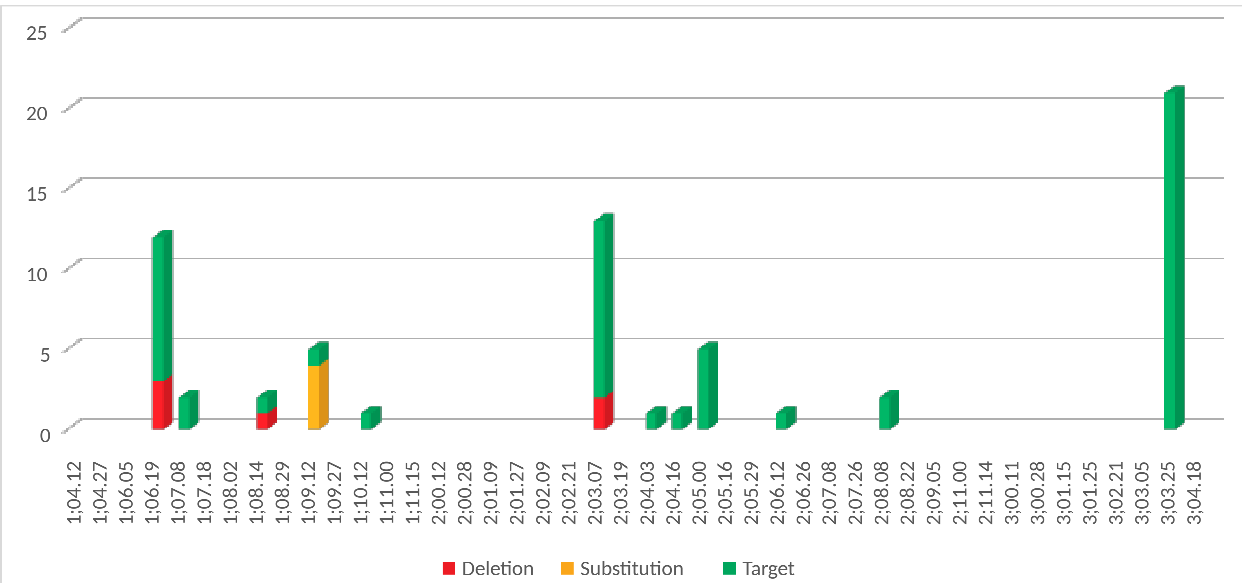
7 This reduction pattern also occurs in complex onsets containing a /w/, but the cluster reduction targets the first consonant rather than the /w/.

onset clusters involving /w/, which will serve as baselines for the description of /Cɹ/ onset clusters.

2.1 Coronal-/w/ onsets

Figure 6 displays William's articulation of /w/ in coronal-/w/ complex onsets across the observation period. Although there are relatively few tokens of this type available from the corpus, most of which are concentrated within a few sessions, we can still see that William had a grasp on the articulation of /w/ in these onsets from the moment they emerged in his productions.

Figure 6: Longitudinal development of coronal-/w/ onset clusters

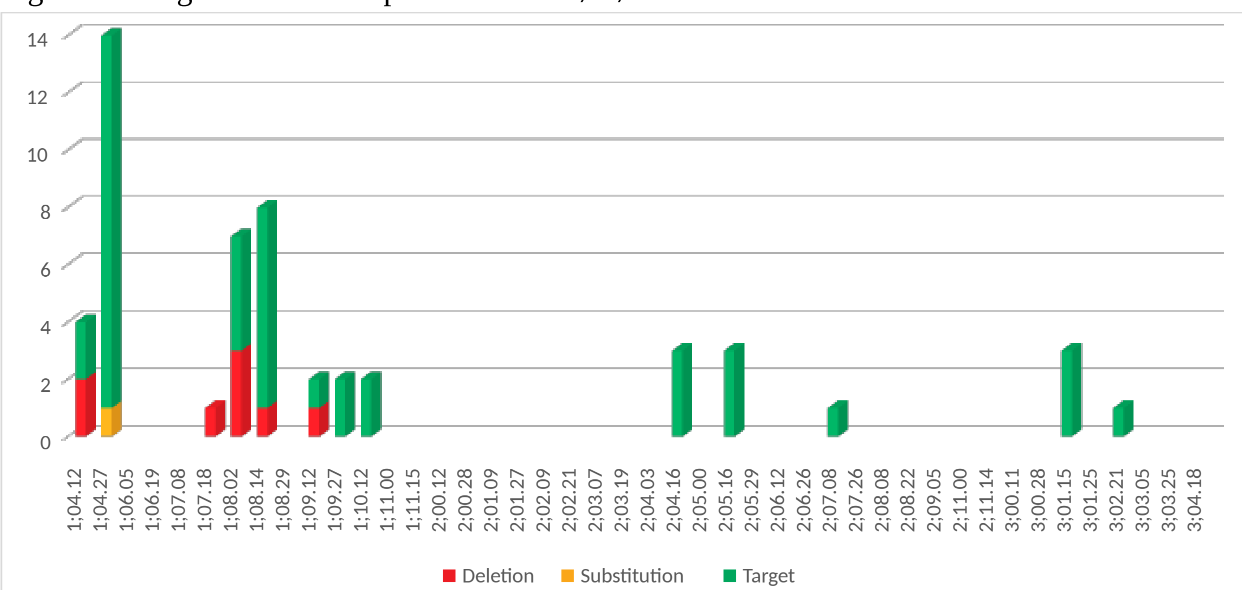


From the first session in which target-like productions of coronal-/w/ onsets are attested, the majority of the /w/s are articulated in a target-like manner. Only in three sessions (1;06.19, 1;09.12, and 2;03.07) do non-target articulations appear in more than one token. The deletions at 1;06.19 and 2;03.07 appear to reflect chance misarticulations in William's speech, as opposed to a systematic pattern of error. Session 1;09.12 displays a large number of substitutions, occurring in an unintelligible word which is assumed to be "swan" by the original transcribers, and which William articulates with an initial [ʒɹ]. These are included due to the best guess for the word documented in the original corpus, although they may not represent true attempts at /sw/ onsets. Aside from these isolated tokens, the overall data suggest that William mastered the articulation of /w/ in coronal-initial complex onsets as early as he did in singleton onsets.

2.2 Velar-/w/ onsets

Figure 7 illustrates William's development of /Cw/ onsets that begin with a velar consonant. Similar to the coronal-initial onsets described above, adult-like articulation of /w/ is prominent from the start of the observation period.

Figure 7: Longitudinal development of velar-/w/ onset clusters

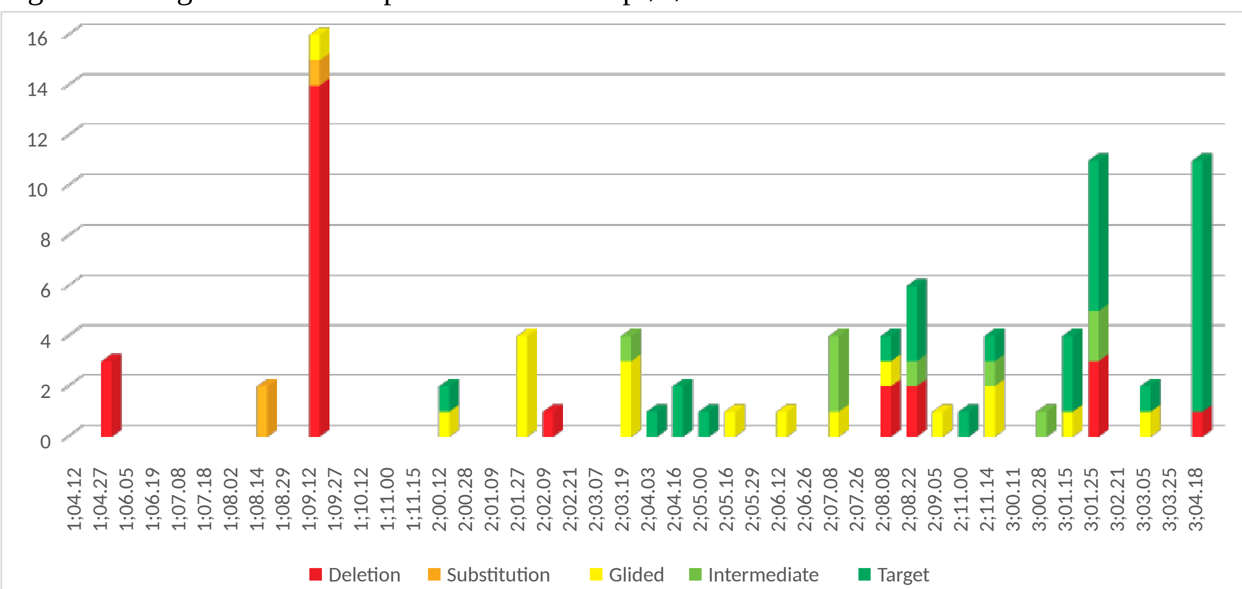


However, we can also observe a number of non-target articulations in sessions prior to 1;09.12, with proportionally more cases of deletion than we observed in coronal-initial clusters. These deletions, particularly in sessions 1;04.12 and 1;08.02, appear in words such as "quack", and are more frequent in utterances in which William produces the word repeatedly in quick succession, and thus may have to do more with undershot articulations than a true reflection of William's productive abilities. Beyond session 1;09.12, each attempt that William makes at velar-/w/ onsets results in a target production of /w/.

2.3 Labial stop-/ɹ/ onsets

As we can see below in Figure 8, William's pronunciation of /ɹ/ in labial stop-/ɹ/ onsets is highly variable throughout the sessions, even past the age at which target-like articulations of /ɹ/ in this context become the majority, at 3;01.15. Prior to this point, deletions and glided or semi-glided articulations variably occur with high frequency, with deletions slightly more prominent in earlier sessions and glided productions representing the majority of William's articulations in later ones.

Figure 8: Longitudinal development of labial stop-/ɹ/ onset clusters

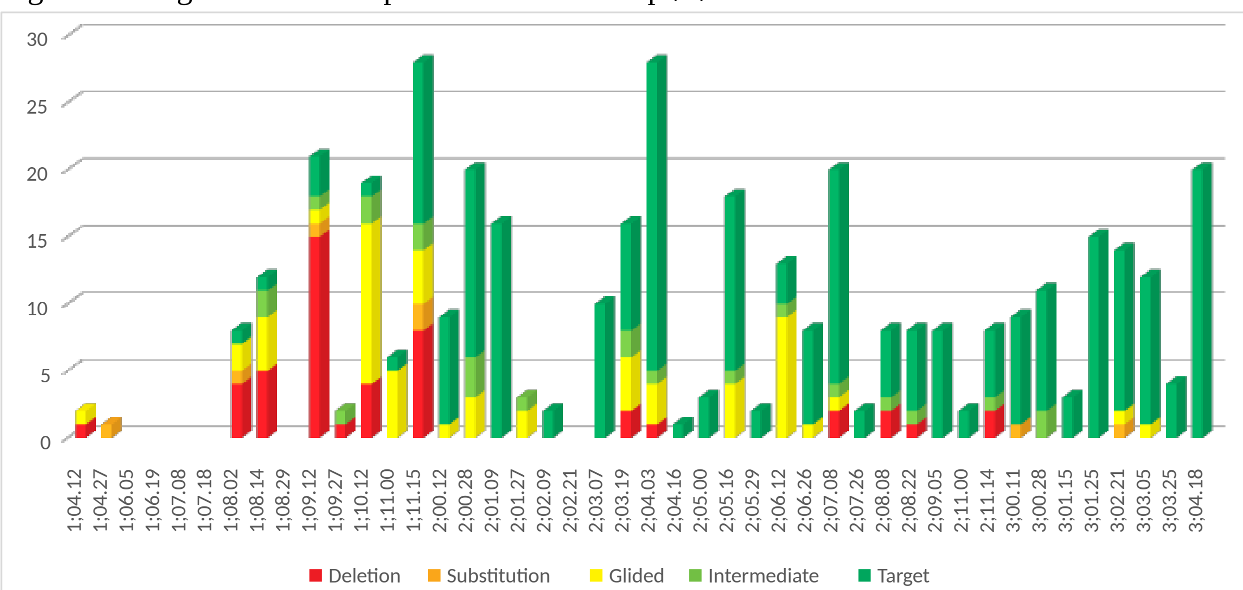


These observations correlate roughly with the period during which /ɹ/ becomes articulated in a majority glided manner in singleton onsets. /ɹ/ becomes target-like in singleton onsets at 2;11.00, however, which is approximately two and a half months before we see a majority of target-like articulations in labial-initial complex onsets. This delay may be an effect of the lower number of tokens for complex labial-/ɹ/ onsets, particularly as we can see target-like articulations starting to appear at 2;11.00. It may also be due to the difficulty for William in articulating the newly-acquired sound as part of a consonant cluster, particularly when the first consonant of the cluster is labial, which may influence the production of the following /ɹ/ to remain more labialized. Aside from these variations, the pattern illustrated by articulations of /ɹ/ in labial-/ɹ/ onsets is very similar to the one observed in singleton /ɹ/ onsets.

2.4 Coronal stop-/ɹ/ onsets

Figure 9 displays the development of coronal stop-/ɹ/ onsets in William's speech. Contrary to the development of /ɹ/ in all other environments, articulations in coronal stop-/ɹ/ onsets reach a stage of apparent mastery around the time that /ɹ/ reaches a stage of apparent gliding or semi-gliding (approximately 1;11.15), both in singleton and all other complex onsets.

Figure 9: Longitudinal development of coronal stop-/ɹ/ onsets

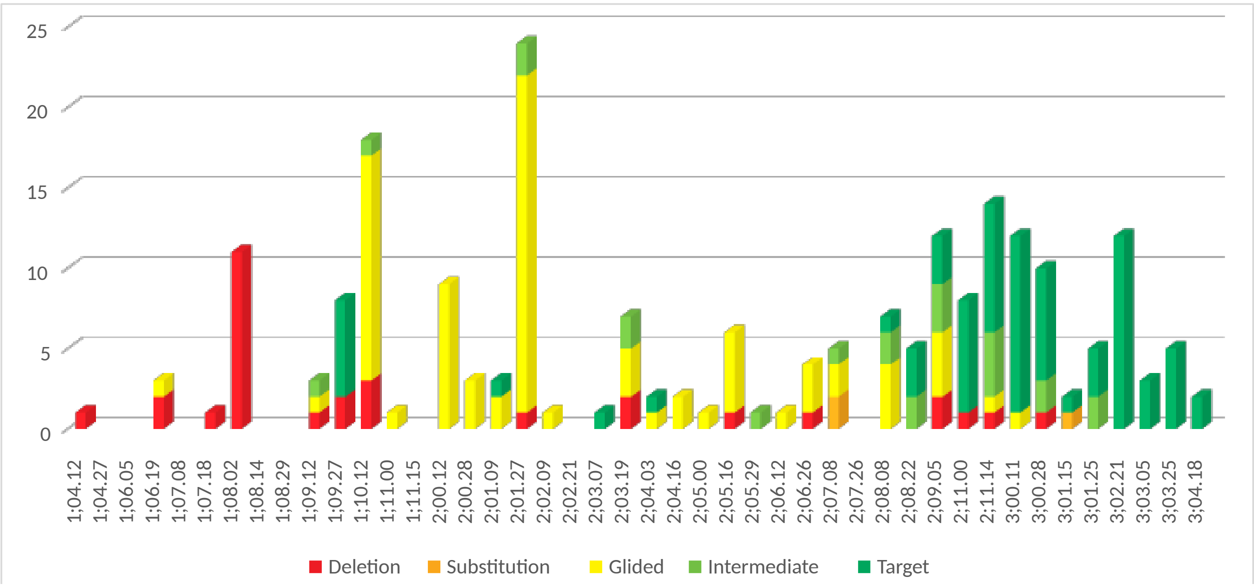


Also of note in this context is a very short period during which William appears to glide /ɹ/ in coronal stop-/ɹ/ onsets, centring on sessions 1;10.12 and 1;11.00. These glided productions occur almost entirely in William's pronunciations of the word "triangle", and therefore may be due simply to lexical influence, as William has an idiosyncratic pronunciation of this word (something like [ʃwaɪŋgo]). However, the most striking feature of coronal-/ɹ/ onsets is the early time of apparent acquisition highlighted above, which raises questions concerning articulatory or phonological influence, or adult listener bias.

2.5 Velar stop-/ɹ/ onsets

As we can see in Figure 10, velar stop-/ɹ/ onsets in William's productions behave similarly to singleton /ɹ/ and labial stop-/ɹ/ onsets with regards to patterns of development. At 1;10.12, /ɹ/ in these complex onsets is transcribed overall as a glide, and this pattern continues until approximately 2;11.00.

Figure 10: Longitudinal development of velar stop-/ɹ/ onset clusters



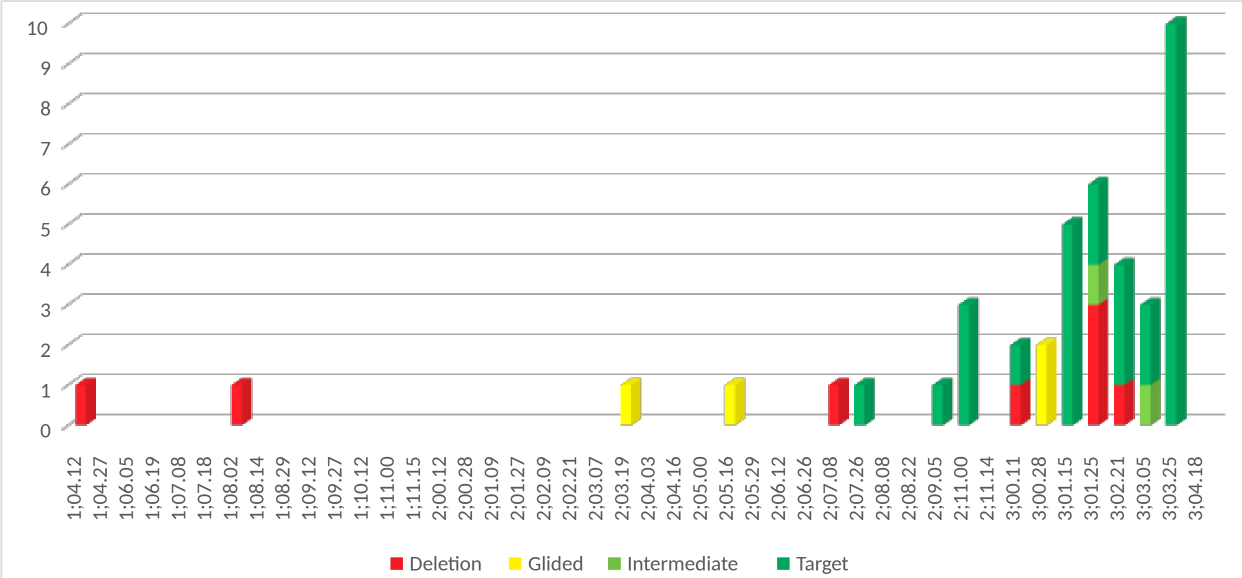
Prior to 1;10.12, variable articulations are common, although William deletes /ɹ/ in most cases. Post-2;11.00, articulations are judged to be correct by adult transcribers. This pattern holds true for almost all cases, with only sessions 1;09.27 and 2;11.14 showing a notable number of articulations which do not fall within these parameters. At 1;09.27, most articulations were transcribed as correct. These target-like pronunciations do not seem to follow any pattern in the data, however. In session 2;11.14, a relatively large number of articulations were transcribed as semi-glided. These articulations occur in words which were also pronounced "correctly" in the same session, and thus may too be chance misarticulations. Aside from this, William's articulations of /ɹ/ patterned quite similarly in velar stop-/ɹ/ onsets as they did in singleton onsets and onset clusters which begin with a labial stop.

2.6 Fricative-/ɹ/ onsets

There were three types of fricative-/ɹ/ onset clusters found in William's speech: /fɹ/, /vɹ/, and /θɹ/. These are illustrated with two graphs and discussed below, with /fɹ/ and /vɹ/ consolidated as labial fricative-/ɹ/, as they display the same pattern, similarly to the voiced and voiceless pairs in stop-initial clusters described just above.

There are very few attempts at words containing labial fricative-/ɹ/ onsets in William's corpus, shown in Figure 11. However, the ones that we do find, for the most part, follow the same patterns as both singleton /ɹ/ onsets and labial/velar stop-/ɹ/ onsets.

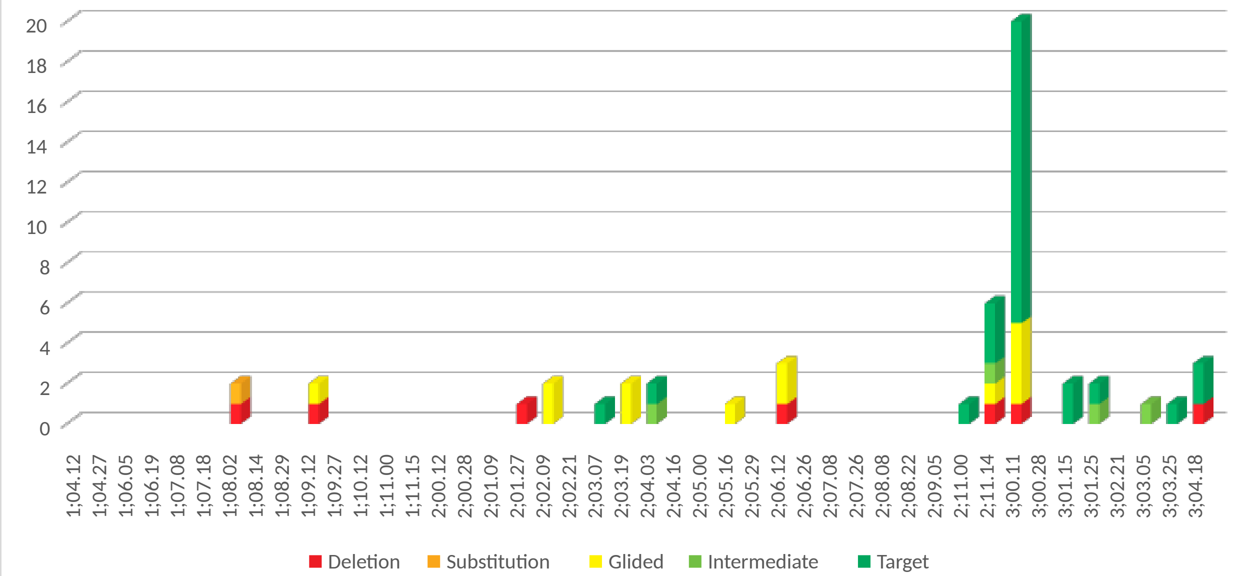
Figure 11: Longitudinal development of labial fricative-/ɹ/ onset clusters



Although there are only two tokens found before 1;10.12, both undergo /ɹ/ deletion. The ones that appear in later sessions prior to the age of mastery are either glided or deleted. Interestingly, target-like articulations appear with a majority in William's productions at 2;07.26, 2;09.05, and 2;11.00. This is earlier than predicted by other environments. However, the low number of tokens in these instances prevents us from drawing any firm conclusions about these data. The overall data suggest that William's articulations of labial fricative-/ɹ/ onsets pattern as expected by singleton /ɹ/ and labial/velar stop-/ɹ/ onsets.

Figure 12 below shows the development of /ɹ/ in William's /θɹ/ onsets. Similar to the labial fricative-/ɹ/ onsets, /θɹ/ onsets mostly display the same patterns as /ɹ/ in singleton onsets. Early articulation is variable, but by the time William reaches two years of age, glided or semi-glided articulations are the norm in his speech; again, a low number of tokens makes this claim tenuous, but this is the general pattern suggested by the data.

Figure 12: Longitudinal development of /θɹ/ onsets



By age 2;11.00, target-like articulations emerge as the dominant production in William's speech. However, this pattern is slightly more variable in the case of /θɹ/ onsets: single instances of target-like articulations occur prior to age 2;11, and non-target like articulations of /ɹ/ are still present in sizable numbers after this time. The variability up until this point illustrates that the benefit in either the articulation or the perception of /ɹ/ seen in section 2.4 for coronal stop-/ɹ/ clusters does not manifest itself in coronal fricative-/ɹ/ onsets. However, note that William produces /θ/ almost invariably as [f], even beyond the age at which he acquires target-like pronunciation of /ɹ/. Thus, we would not expect the 'coronal benefit' observed in the context of /tɹ/ and /dɹ/ clusters to be extended to /ɹ/ in /θɹ/ clusters, as the first consonant is not being produced as a coronal in this latter context. Overall, /θɹ/ articulations in William's speech behave as expected from the patterns displayed by other complex onsets, aside from the coronal stop-initial context.

3 Summary

Taking the impressionistic transcription data as a whole, we see onsets containing /w/ behave in one manner, and onsets containing /ɹ/ behaving uniformly a different way, with one notable exception. William articulates all onsets which contain /w/, both singleton and complex, with a target-like /w/ from the start of the recordings, with only isolated exceptions. For /ɹ/, aside from complex onsets beginning with a coronal stop, William achieves mastery at approximately age 2;11. This statement holds whether the onset is singleton or complex. However, there is a striking difference involving William's complex onsets which begin with a coronal stop. In these onsets, William achieves apparent mastery of /ɹ/ at approximately two years of age, almost one year prior to other onsets containing /ɹ/. This discrepancy raises questions regarding the role of adult interpretation in evaluating child speech, as well as coarticulatory effects on clusters in development. Using formant values taken acoustically from these same speech samples, which I describe in the following chapter, I will return to these questions in Chapter 6.

Chapter 5: Analysis based on acoustic measurements

1 Introduction

In the preceding chapter, I described the development of singleton and complex onsets involving /ɹ/ and /w/ in William's speech over the observation period, as interpreted by transcription data produced by trained linguists. In this chapter, I describe the acoustic measurements obtained from these same speech productions. I compare the data across four contexts: in singleton onsets, I address /ɹ/ productions deemed adult-like from the transcriptions (which I refer to as "rhotic productions"); /ɹ/ productions which have been perceived as [w] ("labialized productions"); and tokens deemed intermediate between an [ɹ] and a [w] ("intermediate productions"). Finally, I compare these three types of /ɹ/ productions against target /w/ tokens, the vast majority of which were judged target-like throughout the observation period ("/w/ productions"), as we saw in the previous chapter. These four patterns and their labels are summarized in the following table. The reason I prefer "labialized" to "glided" from here on will become clearer throughout this and the next chapter; the short version is that there is compelling evidence to analyze this behaviour as a deficit in place of articulation rather than a change in manner of articulation.

Table 5: Behaviours and labels of production patterns of William's /ɹ/ and /w/

Behaviour	Label
/ɹ/→[ɹ]	rhotic productions
/ɹ/→[w]	labialized productions
/ɹ/→[ɹʷ, wʴ]	intermediate productions
/w/→[w]	/w/ productions

Concerning complex onsets, I compare both rhotic and labialized productions of /Cɹ/ onsets against /w/ productions in /Cw/ onsets. (The low number of intermediate productions observed in complex onsets renders those tokens futile to examine in depth.) As we will see in Section 3 of this chapter, however, the data available for complex onsets are sparse and do not offer enough ground to draw firm conclusions, with the exception of coronal stop-initial onsets containing /ɹ/ (i.e. /tɹ/ and /dɹ/ onsets), which display intriguing asymmetries in line with the asymmetric behaviour already noted in Section 2.4 of the preceding chapter, where /ɹ/ in this particular cluster was deemed acquired approximately one year prior to /ɹ/ in any of the other phonological contexts analyzed.

For each production type in singleton and complex onsets, I compare the formant values for F1, F2, and F3, following the method outlined in Chapter 3 (measured as an average of values taken at 50, 60, and 70 percent through the sound for each token), as well as the F3-minus-F2 value, per Chapter 3, Section 3.2.

We can establish two overall stages in William's development of /ɹ/. Before age 2;06, we observe high variability, particularly in the higher formants, and no noticeable differences in the overall patterning of the formants for sounds which are perceived to be similar by adult transcribers. After age 2;06, the formants become more even and regular across the remainder of the recording sessions. The differences between the different productions types also become more apparent, with increasingly clear differences in patterning for the productions perceived as /ɹ/ vs. those perceived as /w/.

We will also witness a further property of the data in coronal stop-initial complex onsets, for which F2 values trend much higher than for the same approximant in either singleton onsets or complex onsets that begin with a non-coronal sound. This disparity raises interesting questions about the possibility of a different kind of contrast between these sounds in William's speech. This issue is discussed further in Chapter 6.

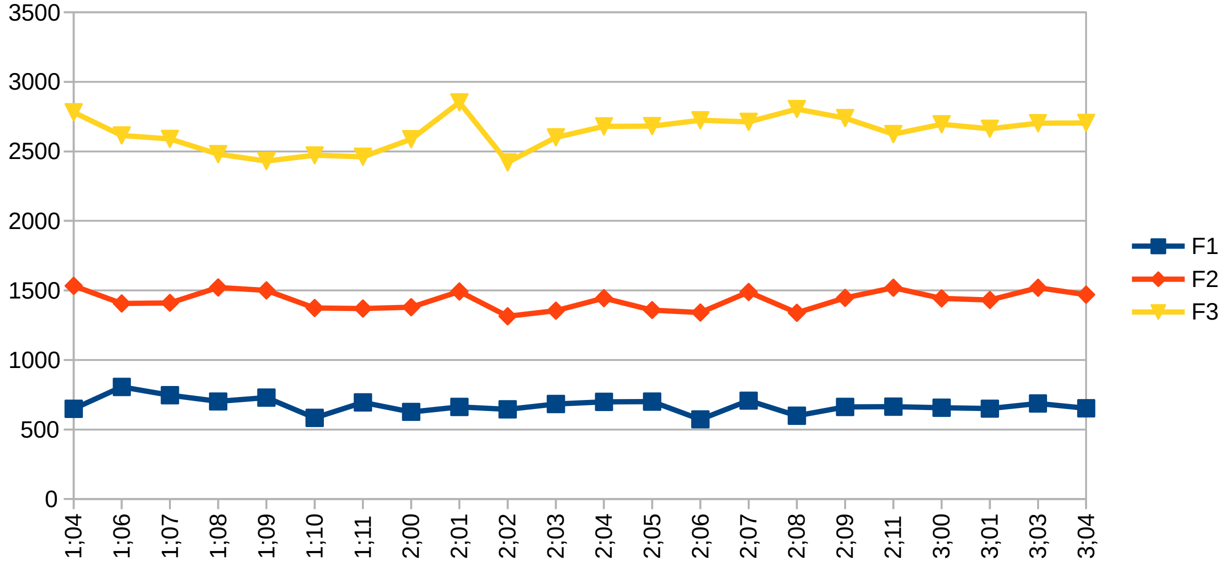
2 Singleton onsets

In this section, I address measurements taken from tokens of /ɹ/ and /w/ in William's singleton onsets. I begin with an analysis of the productions of /w/, followed by the rhotic, labialized, and intermediate productions of /ɹ/. As stated above, the main difference between the formant values measured from these production patterns lies in F2, with minimal differences in F1 and large amounts of variation in F3. Between the two sounds which are perceived as /w/ by adult listeners (i.e. labialized /ɹ/ productions and /w/ productions), F2 and F3-minus-F2 pattern very similarly, which suggests that these sounds were not differentiated in articulation by William, at least not concerning this particular dimension of speech. The formants extracted from /w/ productions differ from those of any type of /ɹ/ production primarily through their relative stability. While /w/ productions display low variability over time, /ɹ/ productions show a higher level of variation, in particular during the period prior to age 2;06.

2.1 /w/ productions

Figure 13 below illustrates the formant measurements for William's productions of singleton /w/ across the monthly time points. Recall that these productions were, on the whole, deemed accurate from the start of the recording sessions, as discussed in Chapter 4, Section 2.1.

Figure 13: Formant measurements of William's /w/ perceptually judged as /w/ (n=1989)

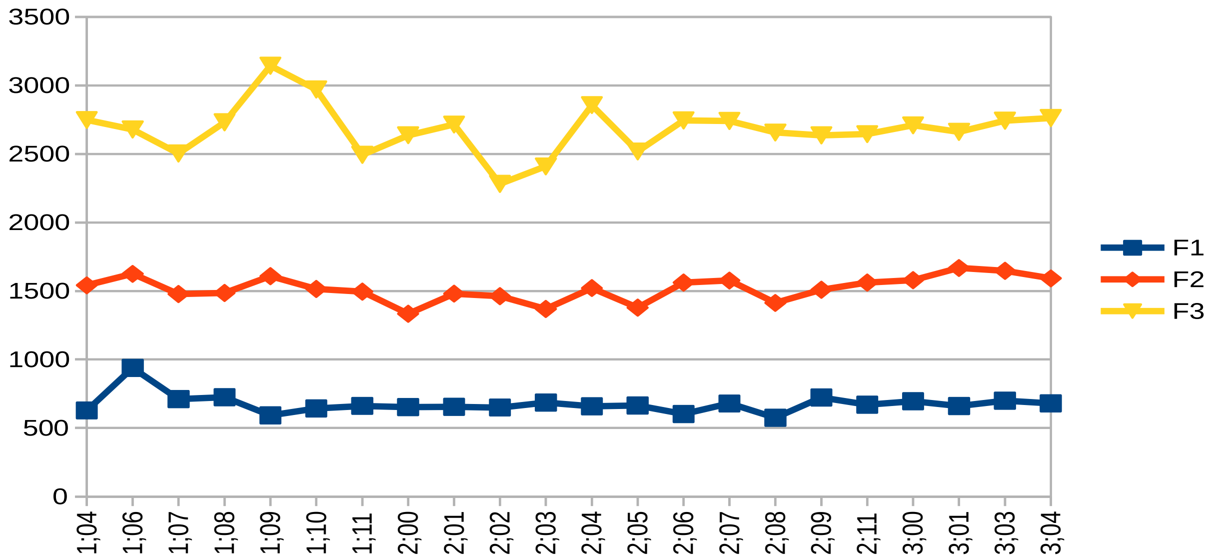


The formant values for William's target productions of /w/ above distinguish themselves from his /ɹ/ productions (reported in Sections 2.2, 2.3, and 2.4) mainly in terms of their relative stability across the observation period, and especially during the earlier part of this period. While we still observe some variability during these earlier sessions – F3 in particular ranging from 2420 to 2852 Hz – the overall trajectory of all values is relatively even across the observation period, for an overall F3-minus-F2 value of 1202-1256 Hz. Finally, the small amounts of variation that we do see can be explained by the general difficulty involved when measuring recordings of child speech. Indeed, Lindblom (1962) posits that formants of child speech can only be measured to a specificity of one-fourth the fundamental frequency, and Buder (1996) points out that differences between child speech formants may often fall below this limit. In light of the fact that William's recordings took place in a natural setting, the level of variability we see within these formant measurements falls well within the range of expected results. More central to the current data description is the fact that the data presented in Figure 13 can provide a baseline estimate against which we can compare and interpret the three types of /ɹ/ productions I describe in the following sections of this chapter.

2.2 Rhotic productions

The chart below illustrates William's formants over time for articulations of /ɹ/ which were judged to be adult-like by adult transcribers.

Figure 14: Formant measurements of William's /ɹ/ perceptually judged as /ɹ/ (n=318)

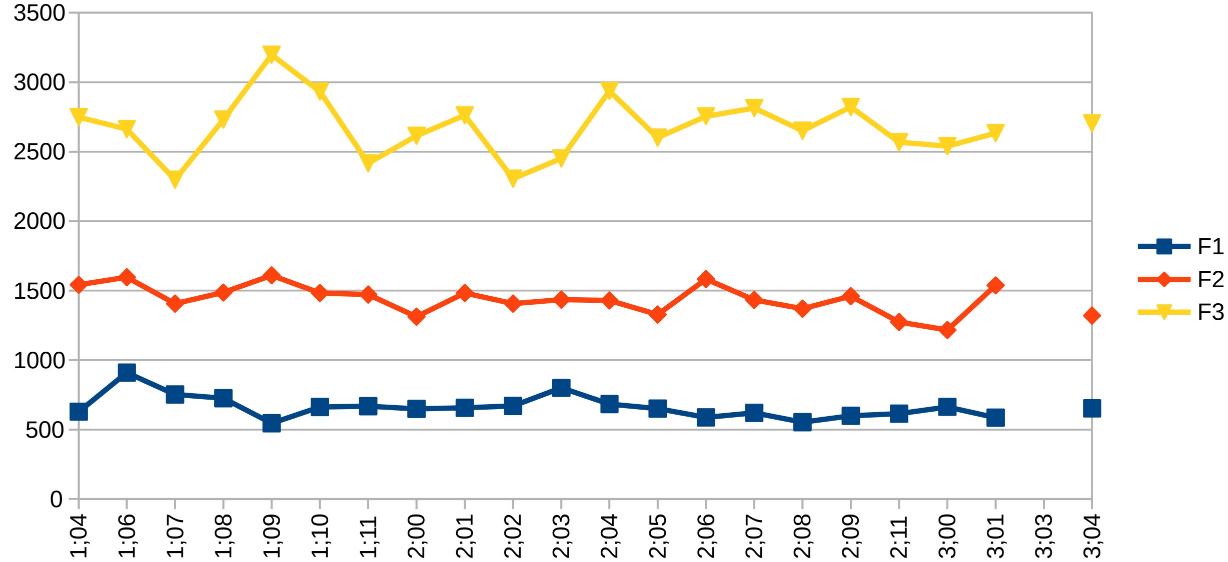


As we can see in this chart, the highest degree of variation is in F3. We can also see F3, as well as F2, become steadier at approximately age 2;06. After that age, F2 generally tends above the 1500 Hz mark, and F3 averages around the 2600-2700 range, giving us an F3-minus-F2 value of approximately 1100-1200 Hz. Note that these formant measurements are taken only from /ɹ/ tokens which have been judged to be target-like, and so some of the variability observed in the earlier sessions may be attributed to the low number of rhotic productions attested before the age of acquisition (n=37, before age 2;06).

2.3 Labialized productions

The following chart illustrates the formant measurements for William's labialized productions of target /ɹ/.

Figure 15: Formant measurements of William's /ɹ/ perceptually judged as /w/ (n=248)



Similar to the graph of rhotic productions in Figure 14, the formants are highly variable up until age 2;06, also with the highest amount of variation affecting F3. After age 2;06, the F3 values stabilize somewhat, although a relatively high level of variation remains. Again, the variation observed may be due to the lower number of labialized productions beyond that age (n=70, compared to 187 before age 2;06).

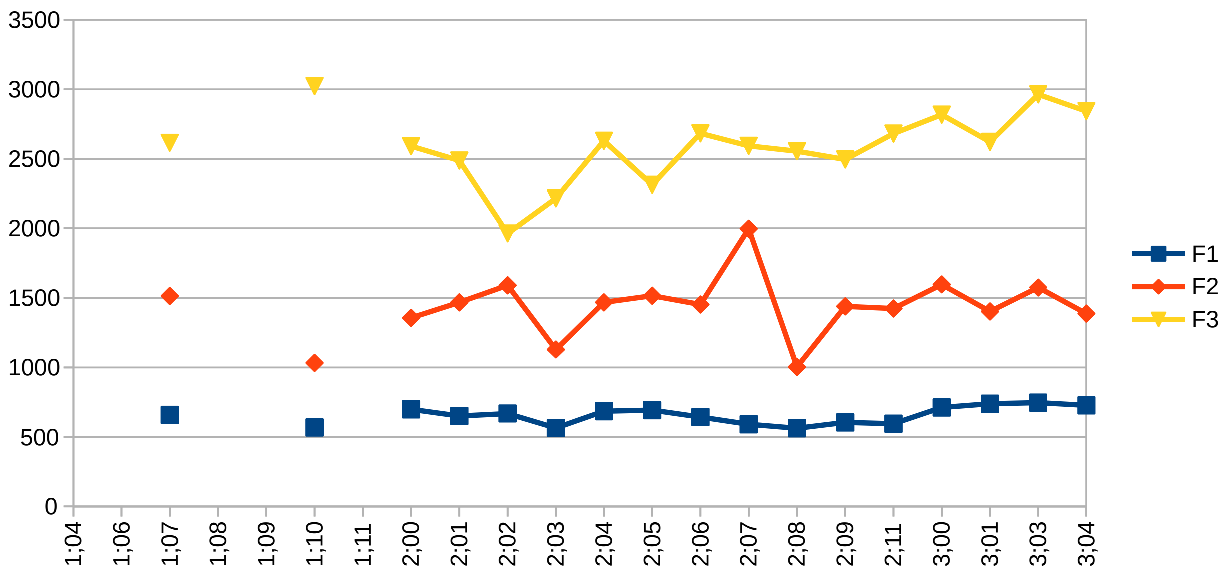
In comparison to Figure 14, the main difference between rhotic and labialized /ɹ/ productions after 2;06 is in the frequency of F2 and the value of F3-minus-F2: the F2 of rhotic productions tends above 1500 Hz, whereas labialized productions have an F2 which falls below the 1500 Hz mark across most of the sessions studied. As the F3 values are similar across the production types, this leads to an F3-minus-F2 value of approximately 1315 Hz for labialized productions of

/ɹ/, vs. a value of around 1058 for rhotic /ɹ/ productions. Recall from Chapter 3, Section 3.1, that a low F3-minus-F2 value is a reliable marker of rhoticity. Finally, the difference between [ɹ]-like and [w]-like productions also reflects itself in the measurements of the target /w/ articulations, which pattern with the labialized /ɹ/ productions through an F3-minus-F2 within the 1202-1256 Hz range, closer to that of labialized /ɹ/ (1315 Hz) than to that of rhotic /ɹ/ productions (1058 Hz).

2.4 Intermediate productions

The following chart depicts the formant values over time in William's productions for target /ɹ/ judged to be intermediate.

Figure 16: Formant measurements of William's /ɹ/ perceptually judged intermediate (n=82)



The number of tokens for intermediate /ɹ/ is lower than for all other types of /ɹ/ productions, especially in earlier sessions (n=43 prior to age 2;06), so any inference we might be able to draw from these data must be taken with a grain of salt. If we look at the mean values for F2 in these data, particularly those occurring in the period after 2;06, they tend between the average F2 values of rhotic and labialized /ɹ/, in line with the general observation that /ɹ/-like and /w/-like productions differ mainly on this parameter.

2.5 Summary

Overall, we observe two distinct phases in William's development of /ɹ/. Prior to age 2;06, we see a high degree of variability in the formant measurements of every type of /ɹ/ production. After 2;06, the data become both more stable and even more measurably different. The most noticeable difference between rhotic and labialized productions of /ɹ/ is in their respective patterning of F2, while the F2 values of intermediate productions generally fall in between these values. The following tables display the average F1, F2, and F3 values for rhotic, labialized, and intermediate productions of /ɹ/ before and after age 2;06, rounded to two decimal points, as well as the F3-minus-F2 values and the total number of tokens for each production.

Table 6: Average F1, F2, F3, and F3-minus-F2 values and number of tokens for rhotic, labialized, and intermediate /ɹ/ before age 2;06

	F1	**F2**	**F3**	**F3-minus-F2**	**n**
Rhotic	649.93	1610.57	2782.20	1171.63	37
Labialized	693.99	1416.77	2644.45	1227.68	187
Intermediate	663.65	1443.70	2437.36	993.66	43

Table 7: Average F1, F2, F3, and F3-minus-F2 values and number of tokens for rhotic, labialized, and intermediate /ɹ/ at/after age 2;06

	F1	**F2**	**F3**	**F3-minus-F2**	**n**
Rhotic	696.37	1649.09	2707.13	1058.04	281
Labialized	606.81	1400.02	2715.31	1315.28	70
Intermediate	668.87	1500.19	2700.02	1199.84	39

We thus observe that /ɹ/ productions transcribed as rhotic display an average F2 value 200-250 Hz higher than that of the labialized productions. As well, the F2 value of the intermediate productions rises over time towards the F2 of the rhotic productions, which might be attributable to William's gradual improvement of his articulation of /ɹ/ as he gets older. The values of F1 and F3 for all three types of production are relatively similar across the tables, with F3 showing more variation prior to age 2;06. We see almost no noticeable difference in F1 across production types, whether before or after age 2;06, with the exception of labialized /ɹ/ productions. After 2;06, the labialized productions display a relatively lower average for F1; again, the number of labialized productions drops significantly after that point, and thus outliers may have a larger effect on the overall data during this developmental period.

On the other hand, we see little variation over time in the values of /w/. The following table depicts the averages of F1, F2, and F3, as well as F3-minus-F2 values and number of /w/ tokens before and after age 2;06.

Table 8: Average F1, F2, F3, and F3-minus-F2 values and number of tokens for /w/ before and at/after age 2;06

	F1	**F2**	**F3**	**F3-minus-F2**	**n**
Pre-2;06	693.75	1408.52	2611.23	1202.71	627
Post-2;06	653.34	1449.65	2706.13	1256.48	1362

In these data, F1 tends between 650 and 700 Hz, which is a very similar range to the F1 values which we saw in Table 6 and Table 7 for all /ɹ/ production types. F2 values are between 1400 and 1450 Hz; this range is similar to the F2 values for labialized productions of /ɹ/, but approximately 200 Hz lower than the F2 values for rhotic productions of /ɹ/. F3 values are between 2600 and 2800 Hz for all production types, except for intermediate productions of /ɹ/

recorded before age 2;06. The F3-minus-F2 values for target /w/ are similar to those of labialized /ɹ/, and both of those values differ from F3-minus-F2 values for target /ɹ/ productions. (Recall from Chapter 3, Section 3.2 that F3-minus-F2 is the most accurate acoustic correlate of rhoticity and the difference between [ɹ] and [w].)

In sum, we see that although the formant values for target /w/ are much smoother across the developmental period than those for labialized /ɹ/, the two sets of values for the production types perceived as [w] (i.e. labialized and /w/ productions) fall within the same ranges overall, but are different from the values of rhotic productions along the parameters of F2 and F3-minus-F2.

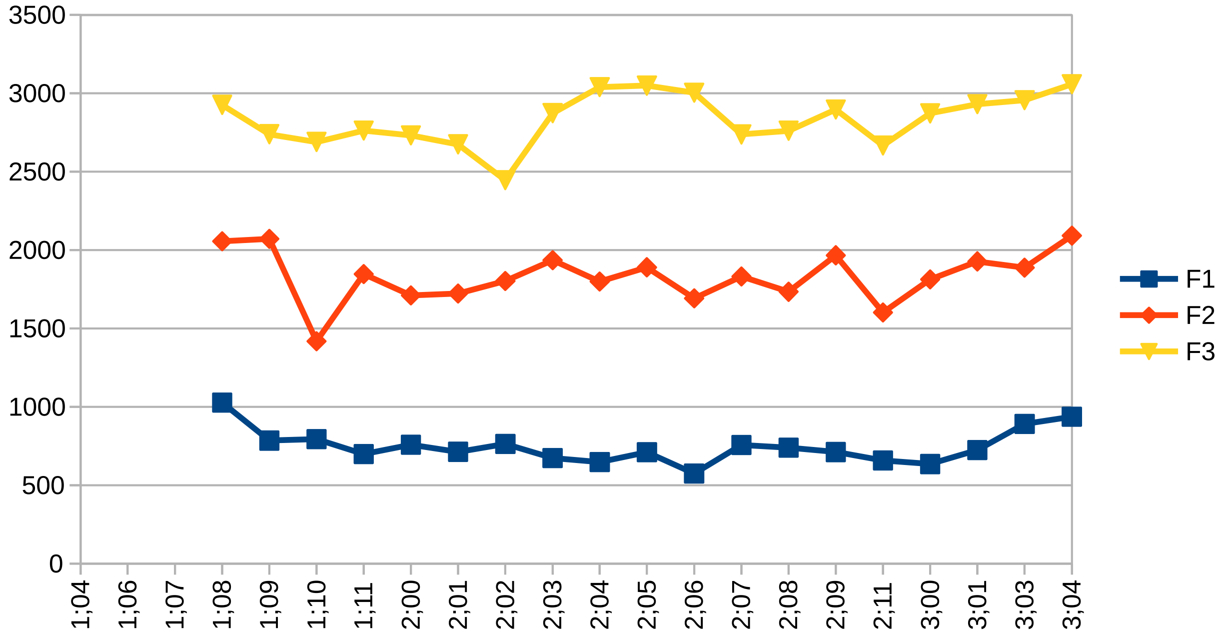
3 Complex onsets

The following section details the development in formant measurements over time for onset clusters containing either /ɹ/ or /w/ in their target forms. As mentioned earlier, relatively little data were available for any particular complex onset environment, with the exception of coronal stop-/ɹ/ onsets. Recall as well that the impressionistic transcription data detailed in Chapter 4, Section 2.4 suggest that /ɹ/ was acquired in coronal stop-initial clusters far earlier than in any other environment. In the acoustic measurements of these clusters discussed below, we observe a similar pattern of development. In addition to coronal-/ɹ/ clusters, I describe the data for coronal-/w/ and velar-/ɹ/ clusters, which provide independent evidence central to our understanding of /ɹ/ in coronal stop-initial onsets. The remainder of the data in complex onsets, which neither comes in large amounts nor provides any extra layers to the current analysis, is grouped together in Section 3.4 for the sake of exhaustivity. I begin this description with coronal-/ɹ/ complex onsets in the next section.

3.1 Coronal-/ɹ/ onsets

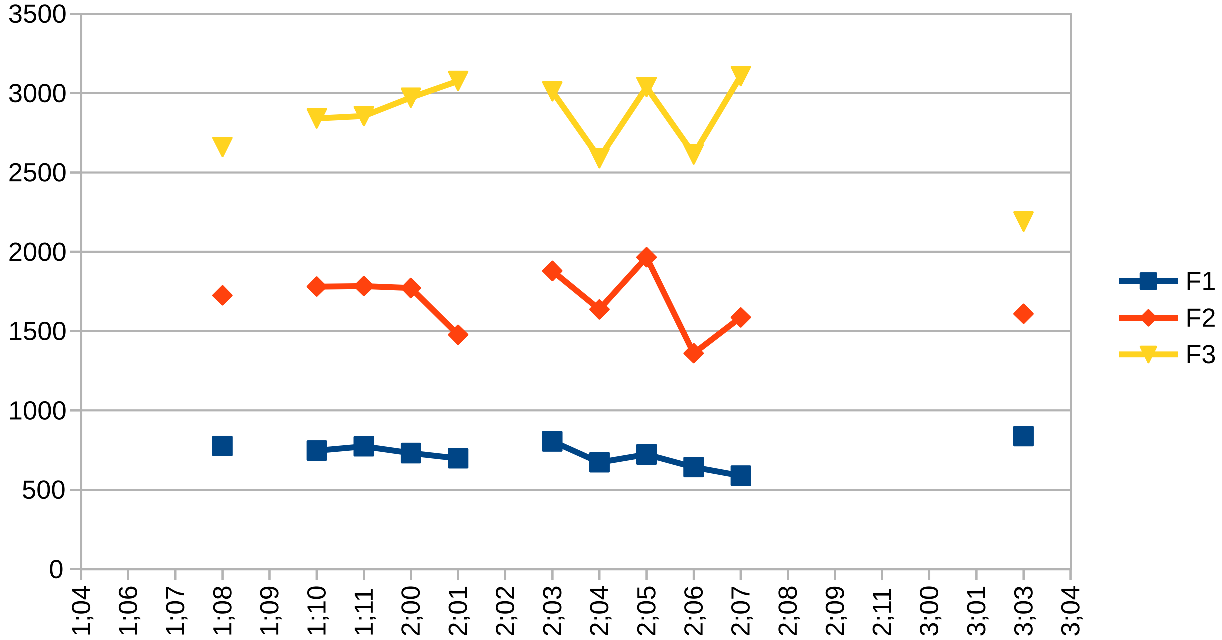
The following graph depicts the formant values for William’s rhotic productions of /ɹ/ in coronal stop-initial complex onsets over time.

Figure 17: Formant measurements of William’s rhotic coronal stop-/ɹ/ clusters over time (n=256)



The number of measurable rhotic tokens in coronal-/ɹ/ onsets is far higher than for any other complex onset in William’s speech. (This context alone provides 256 of the 613 complex onset tokens described in this chapter.) As we can see in Figure 17, the F2 value of /ɹ/ in this environment hovers between 1800 and 2000 Hz, beginning at the time where coronal-/ɹ/ clusters are realized with both target consonants by William (i.e. age 1;11, as discussed in Chapter 4, Section 2.4). The F2 values measured in this context are far higher than those for /ɹ/ in any other environment, whether in singleton or complex onsets. This observation also holds true beyond the age where we observe the emergence of rhotic productions of /ɹ/ in other environments (2;11). This trend is also apparent in the few labialized productions of coronal-/ɹ/ clusters reported in the transcript data, as displayed in Figure 18.

Figure 18: Formant measurements of William's labialized coronal stop-/ɹ/ clusters over time (n=56)

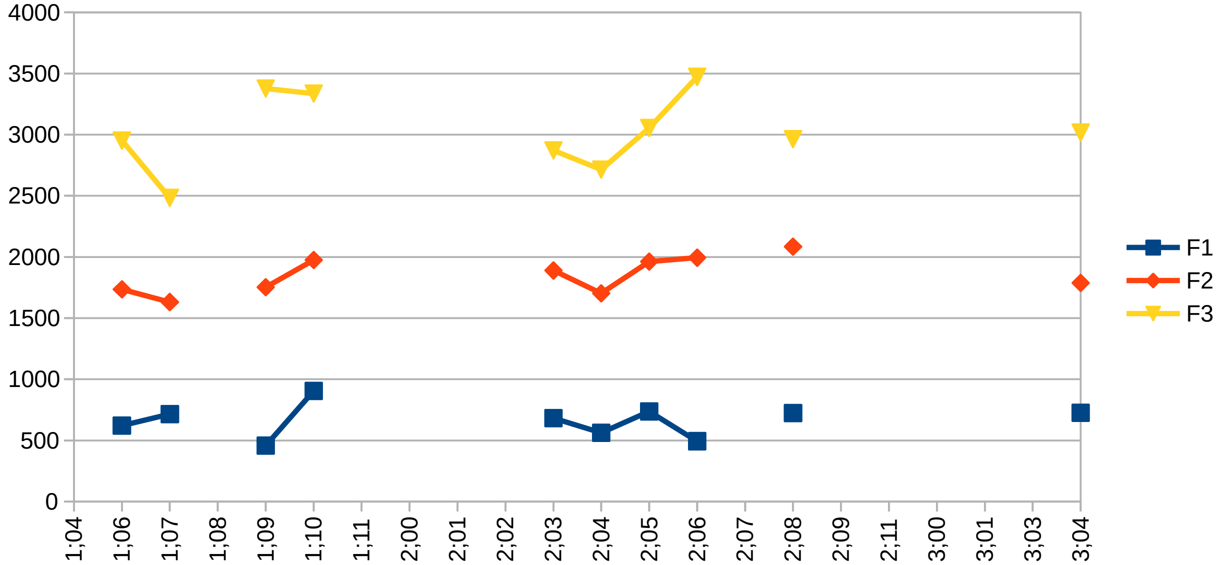


Although we have fewer data points for labialized coronal stop-initial clusters (n=56), the general behaviour of /ɹ/ in coronal-/ɹ/ clusters (i.e. the relative height of F2) is still present in these labialized productions. This uniformity of behaviour suggests some type of context-specific conditioning specific to the cluster-initial coronal stop. I return to this in Chapter 6. In the next section, I compare the formant values of /ɹ/ with those of target /w/ in coronal-initial onsets.

3.2 Coronal-/w/ onsets

The formant measurements over time for William's /w/ in coronal-initial complex onsets are represented in Figure 19.

Figure 19: Formant measurements of William's coronal-/w/ clusters over time (n=55)

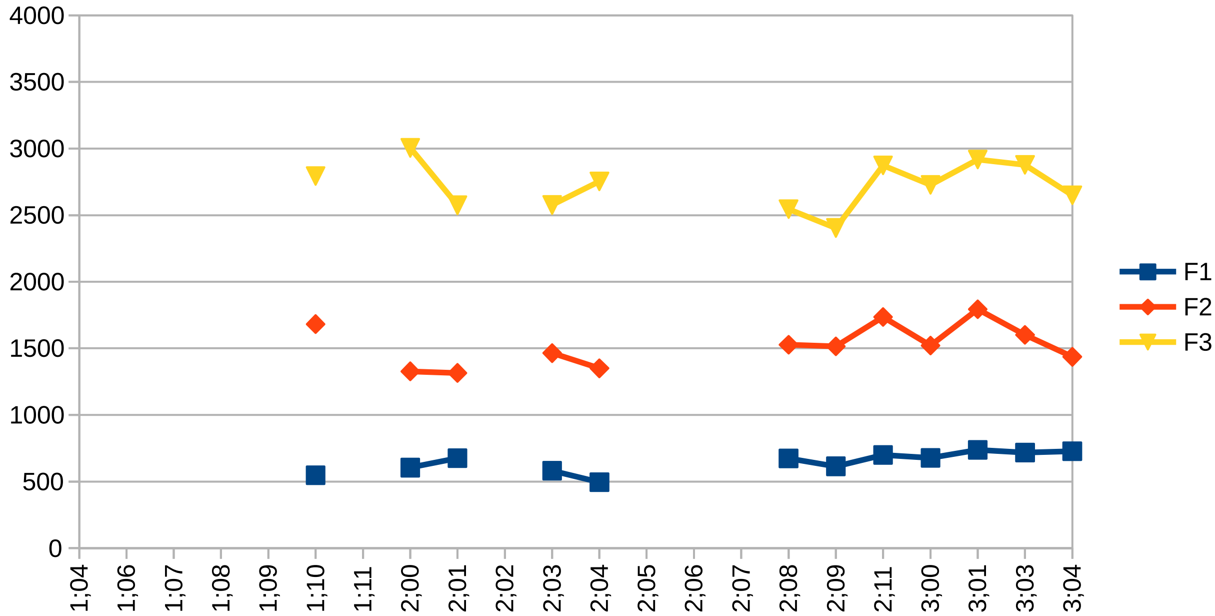


As with many of the other complex onsets (discussed below), the data represented in Figure 19 are sparse (n=55). Additionally, these data are more variable than those for singleton /w/ tokens; the values vary in ways similar to the other onset clusters below. Interestingly, this variability is observed despite William having achieved apparent mastery of singleton /w/ from the beginning of the observation period, further suggesting that the variability in this case is introduced by the presence of another consonant within the onset rather than by anything inherent to /w/. The F2 values for /w/ in these onsets are at times up to 500 Hz higher than those for singleton /w/. As discussed in Section 3.1 above and illustrated in Sections 3.3 and 3.4 below, this high F2 value is specific to complex onsets which begin with a coronal stop. Due to the sparseness and variable nature of the data, however, it is hard to obtain a clear picture of William's coronal-/w/ onsets beyond this raising of F2. Thus, the data presented in Figure 19 are best used as a comparison to the singleton /w/ data of section 2.1 and the coronal-/ɹ/ data of section 3.1 just above.

3.3 Velar-/ɹ/ onsets

The following graphs display formant values for William's rhotic and labialized productions of /ɹ/ in velar-initial onsets.

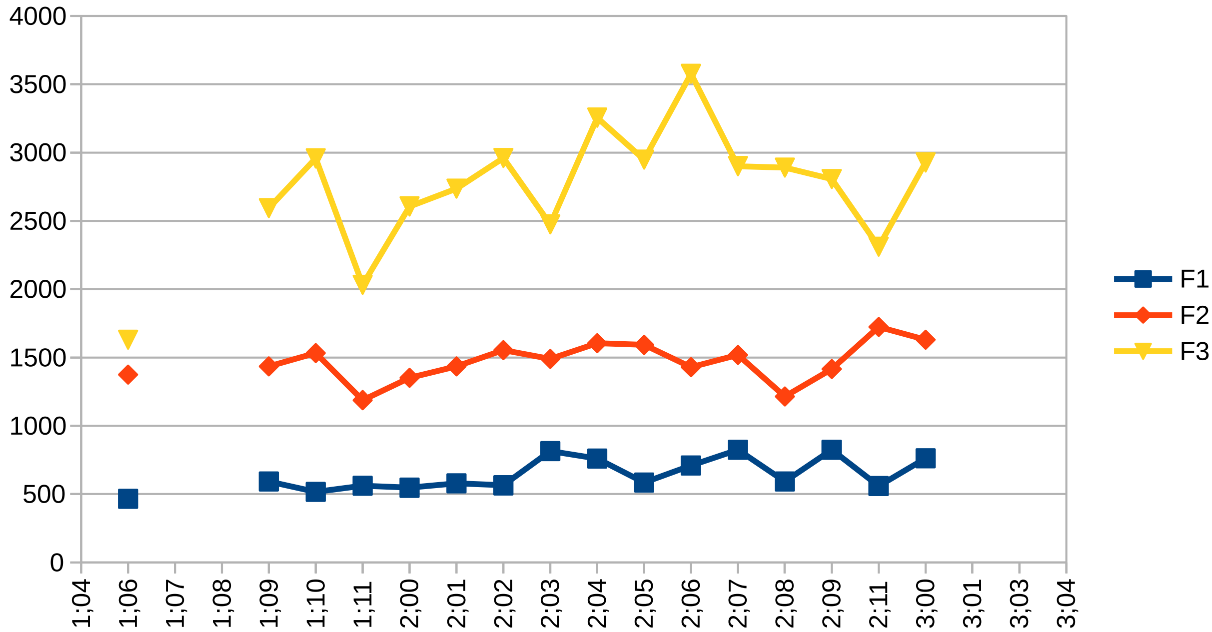
Figure 20: Formant measurements of William's rhotic velar-/ɹ/ clusters over time (n=76)



There are not very many data points to effectively describe the behaviour of rhotic productions of /ɹ/ in velar-/ɹ/ clusters in William's speech (n=76). However, the data are still useful as a contrast to the coronal-/ɹ/ data presented in Figure 17. The key observation is that F2 values in the present context are generally lower than the F2 values observed in coronal-/ɹ/ (and even coronal-/w/) productions.

This lower (i.e. non-raised) F2 value is also seen in the labialized velar-/ɹ/ tokens in Figure 21 below.

Figure 21: Formant measurements of William's labialized velar-/ɹ/ clusters over time (n=82)



The data in Figure 21 provide a strong contrast to the coronal-initial onsets, particularly the coronal-/w/ data in Section 3.2. This holds true in spite of the relatively low number of tokens available for this context (n=82 across the 16 monthly observations), and of the relatively high variability of F3. The data for the labialized tokens are however relatively stable with regards to F1 and F2, with F2 values generally lower than the ones seen in either the coronal-/ɹ/ or the coronal-/w/ productions documented above. This lends credence to the idea that the coronality of the initial consonant in these clusters is causing the F2 of the following approximant to rise.

I describe the remainder of the environments in the next section.

3.4 Residual data

I have grouped together the data for the three remaining contexts in the charts below, due to the low number of tokens available (n=17-42) and the high variation we see in these data across the observation period. They are included here for the sake of completeness, with a brief commentary.

Figure 22: Formant measurements of William’s rhotic labial-/ɹ/ clusters over time (n=29)

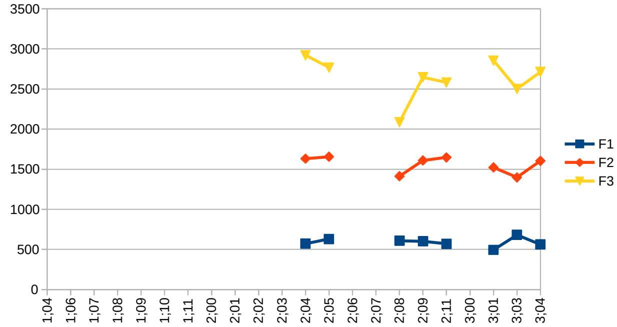


Figure 23: Formant measurements of William’s labialized labial-/ɹ/ clusters over time (n=17)

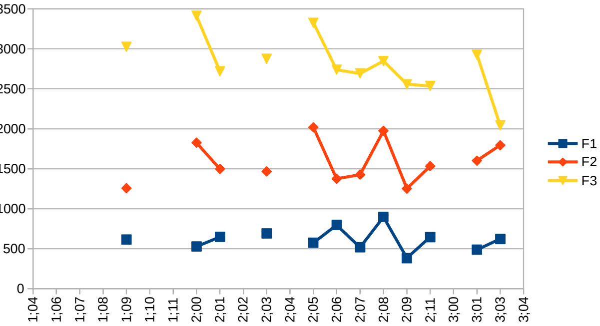
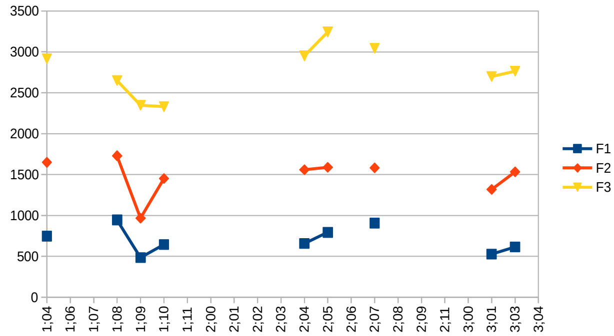


Figure 24: Formant measurements of William’s velar-/w/ clusters over time (n=42)



In these environments, the sparseness and variability of the data are too great to draw any useful conclusions. However, these residual data also do not provide any counter-evidence to the general observations highlighted in the earlier sections. The F2 value for /ɹ/ in non-coronal stop-initial clusters hovers around the values extracted for singleton /ɹ/, which together contrast with the raised F2 values for both /ɹ/ and /w/ observed in coronal-initial clusters.

3.5 Summary

In William's speech, there are relatively few tokens for complex onsets, with the exception of coronal-/ɹ/ onsets. However, the data we do have are quite useful in examining an intriguing asymmetry, as William appears to have acquired correct articulations of coronal-/ɹ/ onsets approximately a year before other /Cɹ/ complex onsets, at least when examined through impressionistic transcriptions (see Chapter 4, Section 2.4). Comparing the acoustic data for coronal-/ɹ/ onsets against those for other environments, we obtain an acoustic correlate to this asymmetry: F2 for /ɹ/ in coronal stop-initial onsets is markedly higher than F2 for /ɹ/ in other complex or singleton onsets. As we have seen in the background literature, as well as through our measurements of /ɹ/ and /w/ in singleton onsets, the distinguishing acoustic correlate of onset /ɹ/ vs. /w/ is the F3-minus-F2 value, with rhotic productions having a lower F3-minus-F2 value than labial(ized) ones. We also notice a marked difference in the height of the F2 value for /w/ in coronal- and velar-initial onsets. Although both were regularly perceived as /w/ in impressionistic analysis, /w/ in coronal-initial onsets also had a higher F2 than the F2 of /w/ in velar-initial or singleton onsets. Taken as a whole, these observations suggest two main points. The first is that an initial coronal stop raises the F2 of a following approximant; this raising of F2, in turn, influences our perception of target /ɹ/. The second point is that this raised F2 cannot be the sole determining factor for our perception of /ɹ/ vs. /w/, due to the fact that /w/ productions following a coronal consonant are still perceived as /w/ despite having a raised F2 when compared to /w/ in other environments. I discuss the role that various factors play in the perception of these two sounds in the following chapter.

Chapter 6: Discussion

In this chapter, I discuss central issues that have emerged through the current analyses, the results of which were described in the preceding two chapters. To begin, I discuss two direct observations of the analyses: (a) the early acquisition of /ɹ/ in coronal stop-initial complex onsets and (b) the lack of any observable distinction between the formant values for labialized /ɹ/ and those for target /w/. Next, I address the implications of these findings for the larger debate centering around the issues of covert contrast and the usefulness of transcription data in the study of acquisition. Additionally, I discuss a slight disparity that we observe between the age at which the formants become more reliably differentiated (2;06) and the time that a large number of target productions arise in the phonetic transcriptions (approximately 2;11). Finally, I include a section outlining the limitations of the current study and conclude with some final remarks.

1 Early mastery of /ɹ/ in coronal stop-initial onsets

Perhaps the most striking observation from the current study is the relative rate at which William acquires /ɹ/ in coronal stop-initial complex onsets in comparison to all other contexts within which /ɹ/ occurs in the data. As suggested by the transcription data, William acquires /ɹ/ in coronal stop-initial onsets approximately one full year before he achieves mastery of the sound in any other environment. There are many potential sources of explanation for this; I discuss a number of them throughout the next subsections.

1.1 F2 values for /ɹ/ and /w/ in coronal-initial onsets

As we saw in Chapter 5, Section 3.1, F2 for /ɹ/ in coronal stop-initial complex onsets in William's speech is approximately 500 Hz higher than that for /ɹ/ in any other environment. As the height of F2 can be correlated with front tongue articulations (Ladefoged & Johnstone 2015), we may argue that the raising of F2 in coronal-/ɹ/ clusters results from a coarticulatory effect triggered by the coronal stop on the following approximant. This rising is in line with previous findings; for example, Hillenbrand, Clark & Nearey (2001) found that the F2 of vowels in adult speech is also raised when the vowel is in contact with a preceding coronal consonant. In support of this hypothesis is the additional observation from this thesis that the F2 of a /w/ in a coronal-initial complex onset is also approximately 500 Hz higher than the F2 for /w/ in other environments. However, given that the value of F3-minus-F2 is the most salient formant feature which distinguishes /ɹ/ from /w/ (Klein et al. 2013), there must be an explanation for why we do not perceive [tɹ] in the case of target /tw/ forms, in spite of the raised F2 affecting the target /w/ within the cluster.

Adult perception of non-standard productions of /u/, the syllabic counterpart to the glide /w/, can also give us insight into how a /w/ with a heightened F2 value might be perceived by English Listeners. Research on dialects such as Californian English, which display widespread patterns of /u/-fronting, has shown that heavily fronted /u/ productions are still perceived as [u] by adult listeners (Koops 2010; Chládková & Hamann 2011). Since English does not have any front or central rounded vowels as part of its phonological (contrastive) inventory, highly fronted /u/ productions may still be parsed by adult listeners as [u] (and thus /w/ as [w]). Another possible explanation for the audible distinction between /tɹ/ and /tw/ that emerges as William starts to produce an adult /ɹ/ in coronal stop-initial clusters, despite the similar F2 values obtained for both /ɹ/ and /w/ in this context, is the relative importance of each formant (F2, F3) for each sound (/ɹ, w/).

In sum, it is plausible, in light of these observations on the perception of fronted /u/ summarized above, that the coronal effect on /w/ productions yields an acoustically distinct allophone that falls below the threshold of perception for adult English listeners. In this respect, we could term this coronalized /w/, with its higher F2 value, a ‘covert allophone’, in that its acoustic differences are not perceptible even to trained listeners (Twist et al. 2007).

1.2 Additional possibilities

Formant values and the differences between F2, F3, and/or F2-minus-F3 values are only a few of several parameters which may distinguish /ɹ/ from /w/ acoustically. We may also consider other possible acoustic (and related articulatory) aspects of speech which may play into our perception of /ɹ/ versus /w/, or the possibility that William was producing a non-target-like contrast (an “inappropriate contrast”, according to Scobbie et al.’s 1996 typology) which was perceived as an appropriate one by the adult transcribers. For example, the amount of lip rounding involved in the production of each sound may play a role in how they are perceived. We could use the relative F3 values of each phone in these tokens to speculate regarding the amount of lip rounding, as F3 lowering has been shown to correlate with this articulatory dimension, with a lower F3 value indicating a higher level of lip protrusion (although F2 height also correlates somewhat with lip rounding as well; West 1999). However, if we wanted to look at the actual degree of lip rounding, rather than simply an acoustic correlate thereof, we would need access to systematic visual recordings of William’s lip configurations during the production of these sounds, recordings which unfortunately do not exist for the current data. Recall that, following a similar line of thinking, I used F2 frequencies as a proxy for tongue positioning in the above discussion. Although several previous studies (as well as the effect of coronal consonants on F2 highlighted in this thesis) suggests that F2 is a relatively robustly acoustic cue for tongue placement, to truly examine tongue shapes and positions in detail would require access to ultrasound imaging for William's speech productions, which we also do not have at our disposal.

More generally, other acoustic parameters may point to measurable differences between /ɹ/ and /w/ as well. Although I took a cursory glance at duration as part of the current study (and found no obvious difference between any of the different production types for /ɹ/ and /w/), I did not examine other parameters such as formant transitions or spectral tilt (Kiefte & Kluender 2005; Kluender & Kiefte 2006), or myriad other aspects of the acoustics which might also show a difference between the production types, also in relation to the relative cues of the following vowel in each context. I leave this exploration to future research.

Another aspect of this study I leave to further research is the actual reason behind William's pattern of /ɹ/ labialization. There are multiple types of explanations as to why this behaviour occurred in William's productions, including articulatory and representational errors. From an articulatory perspective, we could claim that William had some sort of mental representation of /ɹ/ that was distinct from that of /w/. However, a lack of motor control would have caused this difference in representation to be neutralized in articulation (Stevens & Keyser 2010). In contexts where the phonological environment contributed additional articulatory configurations or gestures needed to produce the sound correctly, the difference in representation was carried through the articulation and could be heard overtly. A representational error, on the other hand, would have been caused by a non-adult-like representation in William's phonological lexicon for /ɹ/ (and possibly /w/). This could have taken the shape of a representation of /ɹ/ which was un- or under-specified for the feature [coronal] or [retroflex], resulting in primarily labial productions of this consonant (e.g. Rice 1992; Goad & Rose 2004). However, I leave further discussion of these potential sources of explanation for research on more formal aspects of William's phonological system, which would ideally also incorporate analyses of the other contexts of /ɹ/ production (e.g. syllable codas) in English.

2 Transitional period between developmental stages

As already noted in Chapter 5, the acoustic measurement data suggest a short period during which the formants begin to straighten out for productions such as rhotic and labialized /ɹ/, starting at approximately 2;06. This is approximately five months prior to the time where William begins to produce /ɹ/ as rhotic in a majority of articulations, at age 2;11. (This patterning can be seen in Figure 14 in Chapter 5, Section 3.1.) There are multiple possible explanations for this disparity. One such explanation is that this five-month period involved a a brief stage of actual covert contrast. However, although the formants even out during this time period, they do not display clearly different F2 patterns until approximately around age 2;11. Additionally, during this period, we observe a larger number of correct articulations of /ɹ/ which begin to emerge, although they do not truly become the majority until age 2;11. If we examine the intermediate productions in Figure 16 of Chapter 5, we see the opposite pattern: between ages 2;06 and 2;11, the measurements for F2 are much more variable than either before or after this time. This suggests that this five-month period was transitional, during which William's productions of /ɹ/ gradually improved towards adult-like productions. This hypothesis is in line with the description of the period of covert contrast discussed by Macken & Barton (1980) and Scobbie et al. (1996), in that it involves a period of articulatory "fine-tuning", during which the improvement does not pass the threshold of adult phonological perception. In contrast to their claims, however, this period cannot easily be described as a stage of covert contrast, as the formant values relevant to this period do not present a contrast which was undetectable to adult interpreters. Rather, during this period we see a gradual change in the relative numbers of labialized and target /ɹ/ productions, as can be seen in Figure 4 and Figure 5 of Chapter 4, Section 1. This period, visible in both the transcription data and the acoustic measurements, is thus another example of acoustic and impressionistic data working in tandem.

3 Methodological implications for research on covert contrast and beyond

Regarding the first question I examined from the outset of this thesis research, as to whether covert contrast is a necessary stage in child language development, the data do not show conclusive evidence for the existence of such a stage in William's development of /ɹ/. Indeed, the patterns of contrasts (or lack thereof) perceived and described through the phonetic transcriptions of William’s productions bear out when the same tokens are examined through acoustic analysis. As we saw in Figure 13 and Figure 15 in Chapter 5 above, there is no noticeable difference in the overall formant values for William’s /w/ productions and his labialized /ɹ/ productions. Particularly of note is the lack of difference in F2 (and thus F3-minus-F2) values, along which productions of rhotic and labial(ized) sounds primarily differ both according to previous studies (Espy-Wilson 1992; Borden, Harris & Raphael 1994) and in the data described in Chapter 5 above.

The current study is unique not only in that a longitudinal case study had not yet been performed on the developmental acoustics of the contrast between /ɹ/ and /w/ in English (or, to my knowledge, in any other language); the current study also differs in the way that the tokens were sorted before measurements were taken. Recall that in studies such as Macken & Barton (1980) and Scobbie et al. (1996) summarized in Chapter 2, all tokens of the same target type were analyzed together (e.g. the measurements for all target /st/ productions were compared against all target /t/ productions, irrespective of how they may have been perceived by adult listeners). In contrast to this, instead of pooling together all production types of a single target sound (e.g. rhotic and labialized productions of /ɹ/), I used the transcription data to group each production type into the four categories used in the preceding chapter (i.e. rhotic, labialized, intermediate /ɹ/ productions, and /w/ productions) prior to engaging in acoustic analysis, similar to the methodology used by Dalston (1975) to describe “proper” articulations. Rather than the current study resulting in intermediate formant values, which would have likely been the outcome of pooling all of William’s /ɹ/ data together, it instead revealed relatively systematic parallels between the two types of data: the differences in behaviour noted through IPA transcriptions were clearly reflected in the acoustic data, in particular concerning F2 and F3-minus-F2 values. In light of this, every piece of empirical evidence studied above points to an overt (as opposed to covert) contrast, including both target-like and distorted counterparts of the /ɹ/~/w/ contrast.

Interestingly, on the other hand, we do observe patterns consistent with previous findings on covert contrast in our data. For example, during the stage when William reduced clusters to single consonants (up until approximately 1;09), /tɹ/ and /dɹ/ clusters were not reduced to [t] and [d], but rather to [ʧ] and [ʤ], in line with the coarticulatory pattern of affrication we commonly observe in the production of coronal-/ɹ/ clusters (Vachek 1964). This acoustic trace from the deleted sound is in line with the interpretation by Gulian & Levelt (2011) and Gulian (2017) of the existence of traces of sounds that undergo deletion as a form of covert contrast. In the context of the current discussion, however, these acoustic traces do not form a true covert contrast, as again the transcribers (both the original transcribers and the second transcribers involved in the current study) heard these acoustic traces and transcribed them as such. While the relative accuracy of the phonetic values used in this thesis is no doubt an outcome of the dedicated work of experienced transcribers, similar to the ‘experienced listeners’ in the study by Klein et al. (2012), the fact remains that each measurable production pattern for target /ɹ/ was also perceived impressionistically by the transcribers. Also interesting in this context is the presence of the covert allophone for /w/ in coronal-/w/ clusters, which was clearly measurable acoustically but fell under the same transcribers’ perceptual threshold. Among other questions, this finding calls for additional verifications, both acoustically and perceptually, of this particular phonological context, both in child and adult speech productions.

Note as well that I do not intend, by pointing out the methodological differences highlighted above, to cast aspersions on previous studies such as those by Scobbie et al. (1996) or Richtsmeier (2010); my aim is rather to illustrate that arguments about the potential existence of covert contrasts do not directly translate into arguments against the unsuitability of impressionistic transcriptions for the purposes of examining phonological or phonetic development. The outcomes of the current analysis more centrally raise theoretical questions about the characterization of covert contrast as a *necessary* stage in acquisition.

In the same vein, the literature on phonological and phonetic development provides ample evidence that impressionistic transcriptions alone may not provide a full picture of a child's true productive abilities. This claim is also supported by the current study, at least in one particular context: without acoustic measurements, the covert allophone detected in William's production of coronal-/w/ clusters would have gone unnoticed. A methodology which combines both impressionistic and acoustic analysis is therefore likely to provide the most complete picture of the child's true production abilities.

The current study also highlights potential issues regarding the use of phonetic transcriptions in the current era of data sharing, especially that one should be careful to systematically double-check the transcriptions relevant to the new study, particularly when using data for purposes which were not central to the scope of the original study, as was the case with the current study. Recall from Chapter 3 that the Providence corpus had been, before the current study, primarily used in the area of phonology for studies on syllable coda development (Song, Sundara &

Demuth 2009; Börschinger, Johnson & Demuth 2013). Until the present study, no work had been performed based on this corpus regarding the development of /ɹ/ or /w/ in syllable onsets. When my coworker and I examined the transcriptions and verified the tokens which were relevant to this study, we did make a relatively large number of changes. We can appreciate the differences between the two corpora through a quick comparison of the development of singleton initial onset /ɹ/ in the corpus depicted before our verification and after, as illustrated in the next two figures.

Figure 25: Development of singleton onset /ɹ/, data pre-verification

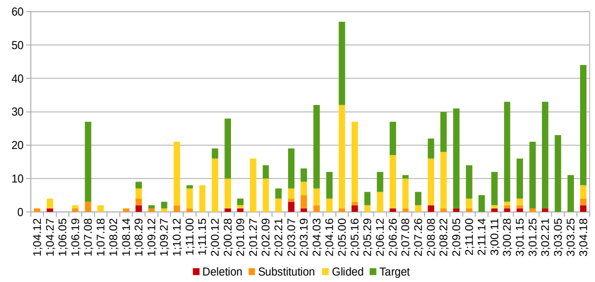
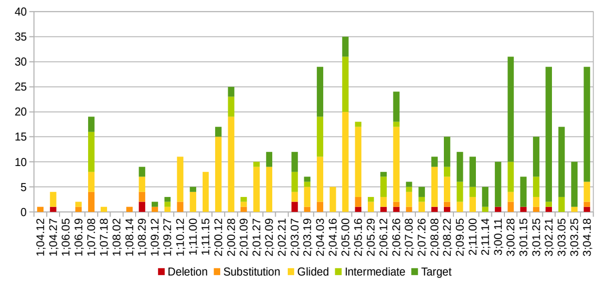


Figure 26: Development of singleton onset /ɹ/, data post-verification



As we see, there is a large difference in the number of tokens found for singleton onset /ɹ/ pre- vs. post-verification. The main reason for this is that a large number of tokens were excluded due to audio quality issues which interfered with either our impressionistic interpretation or the acoustic analysis of the sound (as mentioned in Chapter 3). Were we to include these (excluded) tokens, the numbers would look more similar, but the acoustic analysis of the data would have been compromised. This provides another example of why careful checking of the corpus prior to analysis was an absolutely necessary step.

4 Limitations of the current study

Although I believe the results of this study provide valuable points of discussion for the larger debate regarding covert contrasts and related matters, as discussed in the preceding sections, there are still limitations regarding my choice of method and data to keep in mind. As a case study of a single participant, my study is inherently limited and immediately calls for replication, if only for the sake of independent verification of the patterns observed. William's data were also not perfectly ideal for this study, in a number of ways. As the data were recorded in a naturalistic setting, ambient environmental noise made the recordings less than ideal for acoustic analysis. The naturalistic nature of these recordings also yielded gaps affecting the assessment of certain phonological contexts, which made the interpretation of these contexts more challenging, particularly given the inherent variability observed across the dataset.

In connection to the above, perhaps the biggest limitation of this thesis lies in its lack of statistical analysis, particularly regarding the treatment of formant structures for /ɹ/ and /w/ in Chapter 5. While this is an obvious limitation, I decided to leave this analysis to further work, ideally based upon better quality recordings for clearer analysis, especially the analysis of F3, as the intensity of F3 is a cue to rhoticity of an approximant (Logerquist et al. 2018). My claims concerning the results of the current research (on acoustic values in particular) should thus be taken as a stepping stone towards a more comprehensive study of the development of rhotics in English-learning children, if only to verify the current claims and add more subtlety to the analysis. For example, while I was not able to observe covert contrasts in the data, it is possible that William did display such a a contrast in his development of /ɹ/, if only for a short period of time, something which, again, cannot be fully verified based on the current recordings.

Additionally, as discussed in Section 1 of this chapter, formant measurements are only one way through which a covert contrast may be observed. Other ways through which we might observe a covert contrast include (but are by no means limited to) analyses of formant transitions and articulatory lip movement. As already mentioned, I did take a cursory glance at token duration for each sound production pattern, and also considered formant measurements taken earlier in the sound; they revealed nothing different from those retained for the current study. However, as noted above, it is entirely possible that other acoustic measurements or other empirical approaches such as ultrasound imagery might have provided additional insight on the issues at hand. Taken differently, there is also the fact that in the context where the exhaustive study of all acoustic or articulatory parameters there are is clearly impractical, especially for a longitudinal naturalistic standpoint, the current study actively highlights the usefulness of impressionistic data for studies on the phonetics and phonology of child (and adult) language. In spite of its inherent biases, the human phonological system is clearly efficient at processing all available cues in a dynamic way, as well as being well-optimized to process language-specific speech data. This offers undeniable advantages that remain difficult to match with instrumental analysis.

5 Final remarks

The current case study thus constitutes a challenge to views which claim that covert contrasts represent a necessary stage in child language development. Even though previous studies such as Richtsmeier (2010) have evidenced a covert contrast in the acquisition of /ɹ/ vs. /w/ in English, novel aspects of the method adopted for the current longitudinal acoustic study suggest otherwise, at least in light of the current dataset.

The current work also has implications regarding the suitability of transcriptions for the analysis of speech patterns and child speech development, a widely debated topic in acquisition due to the more fine-grained levels of analysis we can perform with computer-assisted speech analysis, at least on some aspects of the children's speech productions. While transcriptions are inherently limited with regard to our understanding certain aspects of language acquisition, especially in the realm of phonetic development, they are nevertheless an important tool in examining child speech and should not be dismissed altogether. Furthermore, as discussed above, using solely acoustic measurements may introduce other methodological or empirical issues. Perhaps the simplest rejoinder in this case is to adopt the more moderate position that most of the analytic shortcomings inherent to phonetic transcription and acoustic measurement can be addressed when both methods are used in combination.

References

Blake, Frank R. 1917. Reduplication in Tagalog. *The American Journal of Philology* 38(4). 425. doi:10.2307/288967.

Boersma, Paul & David Weenick. 2017. *Praat: Doing Phonetics by Computer*. Amsterdam. http://www.praat.org/.

Borden, Gloria J., Katherine S. Harris & Lawrence J. Raphael. 1994. *Speech Science Primer: Physiology, Acoustics, and Perception of Speech*. Williams & Wilkins.

Börschinger, Benjamin, Mark Johnson & Katherine Demuth. 2013. A joint model of word segmentation and phonological variation for English word-final /t/-deletion. 1508–1516.

Boyce, Suzanne E. & Carol Y. Espy-Wilson. 1997. Coarticulatory stability in American English /r/. *The Journal of the Acoustical Society of America* 101(6). 3741–3753.

Buder, Eugene H. 1996. Experimental phonology with acoustic phonetic methods: formant measures from child speech. In Barbara Bernhardt, John Gilbert & David Ingram (eds.), *Proceedings of the UBC International Conference on Phonological Acquisition* (1), vol. 1, 254–265. University of British Colombia: Cascadilla Press.

Carter, Allyson & LouAnn Gerken. 2004. Do children's omissions leave traces? *Journal of Child Language* 31(3). 561–586.

Chambers, Jack K. (ed.). 2008. *The Handbook of Language Variation and Change* (Blackwell Handbooks in Linguistics). (Blackwell Handbooks in Linguistics). Malden, Mass.: Blackwell.

Chládková, Kateřina & Silke Hamann. 2011. High vowels in Standard British English: /u/-fronting does not result in merger. *Proceedings of the Seventeenth International Conference on Phonetic Sciences*, 476–479. University of Hong Kong.

Churchill, Kimberley. 2009. Covert contrast in the speech of an adolescent with apraxia of speech: a case study. St. John's, Newfoundland: Memorial University of Newfoundland MA Thesis.

Dalston, Rodger M. 1975. Acoustic characteristics of English /w, r, l/ spoken correctly by young children and adults. *The Journal of the Acoustical Society of America* 57(2). 462–467.

Demuth, Katherine, Jennifer Culbertson & Jennifer Alter. 2006. Word-minimality, epenthesis and coda licensing in the early acquisition of English. *Language and Speech* 49(2). 137–173. doi:10.1177/00238309060490020201.

Disner, Sandra Ferrari. 1980. Evaluation of vowel normalization procedures. *The Journal of the Acoustical Society of America* 67(1). 253–261. doi:10.1121/1.383734.

Espy-Wilson, Carol 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.

Espy-Wilson, Carol Y., Suzanne E. Boyce, Michel Jackson, Shrikanth Narayanan & Abeer Alwan. 2000. Acoustic modeling of American English /r/. *The Journal of the Acoustical Society of America* 108(1). 343–356. doi:10.1121/1.429469.

Evans, Karen E. & Katherine Demuth. 2012. Individual differences in pronoun reversal: Evidence from two longitudinal case studies. *Journal of Child Language* 39(01). 162–191. doi:10.1017/S0305000911000043.

Ferguson, Charles A. 1964. Baby Talk in Six Languages. *American Anthropologist* 66. 103–114. doi:10.1525/aa.1964.66.suppl_3.02a00060.

Fikkert, P. 1994. On the acquisition of prosodic structure. Leiden: Rijksuniversiteit Leiden PhD Dissertation. https://repository.ubn.ru.nl/handle/2066/32125 (19 September, 2019).

Forrest, Karen, Gary Weismer, Megan Hodge, Daniel A. Dinnsen & Mary Elbert. 1990. Statistical analysis of word-initial /k/ and /t/ produced by normal and phonologically disordered children. *Clinical Linguistics & Phonetics* 4(4). 327–340. doi:10.3109/02699209008985495.

Goad, Heather & Yvan Rose. 2004. Input elaboration, head faithfulness, and evidence for representation in the acquisition of left-edge clusters in West Germanic. In René Kager, Joe Pater & Wim Zonneveld (eds.), *Constraints in Phonological Acquisition*, 109–157. Cambridge: Cambridge University Press. doi:10.1017/CBO9780511486418.005.

Gulian, Margarita. 2017. The development of the speech production mechanism in young children: Evidence from the acquisition of onset clusters in Dutch. Utrecht, The Netherlands: Netherlands Graduate School of Linguistics PhD Dissertation.

Gulian, Margarita & Claartje Levelt. 2009. An acoustic analysis of child language productions with reduced clusters. In Jane Chandlee, Michelle Franchini, Sandy Lord & Marion Rheiner (eds.), *BUCLD 33 Online Proceedings Supplement*, 1–12. Somerville, MA: Cascadilla Press.

Gulian, Margarita & Claartje Levelt. 2011. Temporal measures of reduced /sC/-clusters in toddler speech: evidence for a detailed lexical specification. *Proceedings of the Seventeenth International Congress on Phonetic Sciences*, 787–790. University of Hong Kong.

Hale, Mark & Charles Reiss. 1998. Formal and empirical arguments concerning phonological acquisition. *Linguistic Inquiry* 29(4). 656–683. doi:10.1162/002438998553914.

Hansson, Gunnar Ólafur. 2010. *Consonant Harmony: Long-Distance Interaction in Phonology*. Berkeley: University of California Press.

Hewlett, Nigel. 1988. Acoustic properties of /k/ and /t/ in normal and phonologically disordered speech. *Clinical Linguistics & Phonetics* 2(1). 29–45. doi:10.3109/02699208808985242.

Hillenbrand, J. M., M. J. Clark & T. M. Nearey. 2001. Effects of consonant environment on vowel formant patterns. *The Journal of the Acoustical Society of America* 109(2). 748–763. doi:10.1121/1.1337959.

Ingram, David. 1974. Fronting in child phonology. *Journal of Child Language* 1(2). 49–64. doi:10.1017/S0305000900000672.

Inkelas, Sharon & Yvan Rose. 2007. Positional neutralization: a case study from child language. *Language* 83(4). 707–736.

Jakobson, Roman. 1941. *Kindersprache, Aphasie, und allgemeine Lautgesetze*. Uppsala: Almqvist & Wiksell.

Kiefte, Michael & Keith R. Kluender. 2005. The relative importance of spectral tilt in monophthongs and diphthongs. *The Journal of the Acoustical Society of America* 117(3). 1395–1404. doi:10.1121/1.1861158.

Klein, Harriet B., Lisa Davidson & Maria I. Grigos. 2009. Relationship between perceptual and ultrasound measurements of /r/ development in children with phonological delay. Poster. Paper presented at the Annual Convention of the International Child Phonology Group, Austin, TX.

Klein, Harriet B., Maria I. Grigos, Tara McAllister Byun & Lisa Davidson. 2012. The relationship between inexperienced listeners' perceptions and acoustic correlates of children's /r/ productions. *Clinical Linguistics & Phonetics* 26(7). 628–645. doi:10.3109/02699206.2012.682695.

Klein, Harriet B., Tara McAllister Byun, Lisa Davidson & Maria I. Grigos. 2013. A multidimensional investigation of children's /r/ productions: perceptual, ultrasound, and acoustic measures. *American Journal of Speech-Language Pathology* 22(3). 540. doi:10.1044/1058-0360.

Kluender, Keith R. & Michael Kiefte. 2006. Speech Perception within a Biologically Realistic Information-Theoretic Framework. *Handbook of Psycholinguistics*, 153–199. Elsevier. doi:10.1016/B978-012369374-7/50007-9. https://linkinghub.elsevier.com/retrieve/pii/B9780123693747500079.

Koops, Christian. 2010. /u/-Fronting is not Monolithic: Two Types of Fronted /u/ in Houston Anglos. *University of Pennsylvania Working Papers in Linguistics* 16(2). https://repository.upenn.edu/pwpl/vol16/iss2/14.

Kornfeld, Judy R. & Henry Goehl. 1974. A new twist to an old observation: kids know more than they say. In Anthony Bruck, Robert A. Fox & Michael W. La Galy (eds.), *Papers from the Chicago Linguistic Society Parasession on Natural Phonology*, 210–219. Chicago, Illinois: Chicago Linguistic Society.

Ladefoged, Peter & Keith Johnstone. 2015. *A Course in Phonetics*. Seventh edition. Stamford, CT: Cengage Learning.

Lin, Susan & Katherine Demuth. 2013. The gradual acquisition of English /l/. In Sarah Baiz, Nora Goldman & Rachel Hawkes (eds.), *Boston University Conference on Language Development Proceedings* (37), 206–218. Boston University: Cascadilla Press.

Lindblom, Bjorn. 1962. Accuracy and limitations of sonagraph measurements. In Antti Sovijärvi & Pentti Aalton (eds.), *Proceedings of the Fourth International Congress of Phonetic Sciences*, vol. 1, 188–202. The Hague: Mouton.

Logerquist, Mara, Hyuna A. Kim, Alisha B. Martell, Benjamin Munson & Jan Edwards. 2018. Growth in the accuracy of preschool children's /r/ production: Evidence from a Longitudinal Study. *The Journal of the Acoustical Society of America* 143(3). 1970–1970. doi:10.1121/1.5036492.

Macken, Marlys A. & David Barton. 1980. The acquisition of the voicing contrast in English: a study of voice onset time in word-initial stop consonants. *Journal of Child Language* 7(1). 41–74. doi:10.1017/S0305000900007029.

McAllister Byun, Tara, Adam Buchwald & Ai Mizoguchi. 2016. Covert contrast in velar fronting: an acoustic and ultrasound study. *Clinical Linguistics & Phonetics* 30(3–5). 249–276. doi:10.3109/02699206.2015.1056884.

McAllister Byun, Tara, Sharon Inkelas & Yvan Rose. 2016. The A-map model: articulatory reliability in child-specific phonology. *Language* 92(1). 141–178. doi:10.1353/lan.2016.0000.

McGowan, Richard S., Susan Nittrouer & Carol J. Manning. 2004. Development of [ɹ] in young, Midwestern, American children. *The Journal of the Acoustical Society of America* 115(2). 871–884.

Munson, Benjamin, Jan Edwards, Sarah K. Schellinger, Mary E. Beckman & Marie K. Meyer. 2010. Deconstructing phonetic transcription: covert contrast, perceptual bias, and an extraterrestrial view of vox humana. *Clinical Linguistics & Phonetics* 24(4–5). 245–260. doi:10.3109/02699200903532524.

Pasch, Helma. 2017. Verbal plural in Zande. *Sprachtypologie und Universalienforschung; Berlin* 70(1). 215.

Priestly, Tom M. S. 1977. One idiosyncratic strategy in the acquisition of phonology. *Journal of Child Language* 4(1). 45–65. doi:10.1017/S0305000900000477.

Rice, Keren D. 1992. On deriving sonority: a structural account of sonority relationships. *Phonology* 9(1). 61–99. doi:10.1017/S0952675700001500.

Richtsmeier, Peter. 2010. Child phoneme errors are not substitutions. *Toronto Working Papers in Linguistics* 33. 1–15.

Rose, Yvan. 2000. Headedness and prosodic licensing in the L1 acquisition of phonology. McGill University PhD Dissertation.

Rose, Yvan. 2017. Child phonology. *Oxford Research Encyclopedia of Linguistics*. doi:10.1093/acrefore/9780199384655.013.150.

Rose, Yvan & Katherine Demuth. 2006. Vowel epenthesis in loanword adaptation: representational and phonetic considerations. *Lingua* 116(7). 1112–1139. doi:10.1016/j.lingua.2005.06.011.

Rose, Yvan & Sharon Inkelas. 2011. The interpretation of phonological patterns in first language acquisition. In Colin J. Ewen, Elizabeth Hume, Marc van Oostendorp & Keren Rice (eds.), *The Blackwell Companion to Phonology*, 2414–2438. Malden, MA: Miley-Blackwell.

Rose, Yvan & Brian MacWhinney. 2014. The PhonBank project: data and software-assisted methods for the study of phonology and phonological development. In Jacques Durand, Ulrike Gut & Gjert Kristofferson (eds.), *The Oxford Handbook of Corpus Phonology*, 308–401. Oxford: Oxford University Press.

Rose, Yvan, Brian MacWhinney, Rodrigue Byrne, Gregory Hedlund, Keith Maddocks, Philip O'Brien & Todd Wareham. 2006. Introducing Phon: a software solution for the study of phonological acquisition. In David Bamman, Tatiana Magnitskaia & Colleen Zaller (eds.), *Proceedings of the 30th Annual Boston University Conference on Language Development*, 489–500. Somerville, MA: Cascadilla Press.

Scobbie, James M., Fiona Gibbon, William J. Hardcastle & Paul Fletcher. 1996. Covert contrast as a stage in the acquisition of phonetics and phonology. (Ed.) James M. Scobbie. *QMC Working Papers in Speech and Language Sciences* 1. 43–62.

Shoji, Shinichi & Kazuko Shoji. 2014. Vowel epenthesis and consonant deletion in Japanese loanwords from English. *Proceedings of the Annual Meetings on Phonology* 1(1). doi:10.3765/amp.v1i1.16.

Smit, Ann B. 1993. Phonologic error distributions in the Iowa-Nebraska Articulation Norms Project: consonant singletons. *Journal of Speech and Hearing Research* 36(3). 533–547.

Smith, N. V. 1973. *The Acquisition of Phonology: A Case Study*. Cambridge: Cambridge University Press.

Song, Jae Yung, Megha Sundara & Katherine Demuth. 2009. Phonological constraints on children's production of English third person singular -s. *Journal of Speech, Language, and Hearing Research* 52(3). 623–642. doi:10.1044/1092-4388(2008/07-0258).

Stevens, Kenneth Noble & Samuel Jay Keyser. 2010. Quantal theory, enhancement and overlap. *Journal of Phonetics* 38(1). 10–19. doi:10.1016/j.wocn.2008.10.004.

Stokes, Stephanie F. & Valter Ciocca. 1999. The substitution of [s] for aspirated targets: perceptual and acoustic evidence from Cantonese. *Clinical Linguistics & Phonetics* 13(3). 183–197. doi:10.1080/026992099299130.

Tryon, Darrell T. 1995. *Comparative Austronesian Dictionary: An Introduction to Austronesian Studies*. Berlin: Mouton de Gruyter.

Twist, Alina, Adam Baker, Jeff Mielke & Diana Archangeli. 2007. Are 'covert' /ɹ/ allophones really indistinguishable? *University of Pennsylvania Working Papers in Linguistics* 13(2). 207–216.

Tyler, Ann A., Mary Louise Edwards & John H. Saxman. 1990. Acoustic validation of phonological knowledge and its relationship to treatment. *The Journal of Speech and Hearing Disorders* 55(2). 251–261.

Vachek, Josef. 1964. On peripheral phonemes of modern English. *Brno Studies in English*, vol. 4, 7–109. Prague, Czech Republic: Statni Pedagogicke Nakladatelstvi.

Watters, David E. 2009. *A Grammar of Kham*. Cambridge: Cambridge University Press.

West, Paula. 1999. The extent of coarticulation of English liquids: An acoustic and articulatory study. *Proceedings of the Fourteenth International Congress of Phonetic Sciences*, 1901–1904. San Francisco, United States: Cambridge University Press.

Young, Edna C. & Harvey R. Gilbert. 1988. An analysis of stops produced by normal children and children who exhibit velar fronting. *Journal of Phonetics* 16(2). 243–246.

Young, Robert W. & William Morgan. 1987. *The Navajo Language: A Grammar and Colloquial Dictionary*. Albuquerque: University of New Mexico Press.

Appendix A: Graphs of William's development of /ɹ/ and /w/ pronunciation divided by stress, position, and initial consonant of complex onset

Figure 27: Longitudinal development of word-initial stressed /w/ in singleton onsets

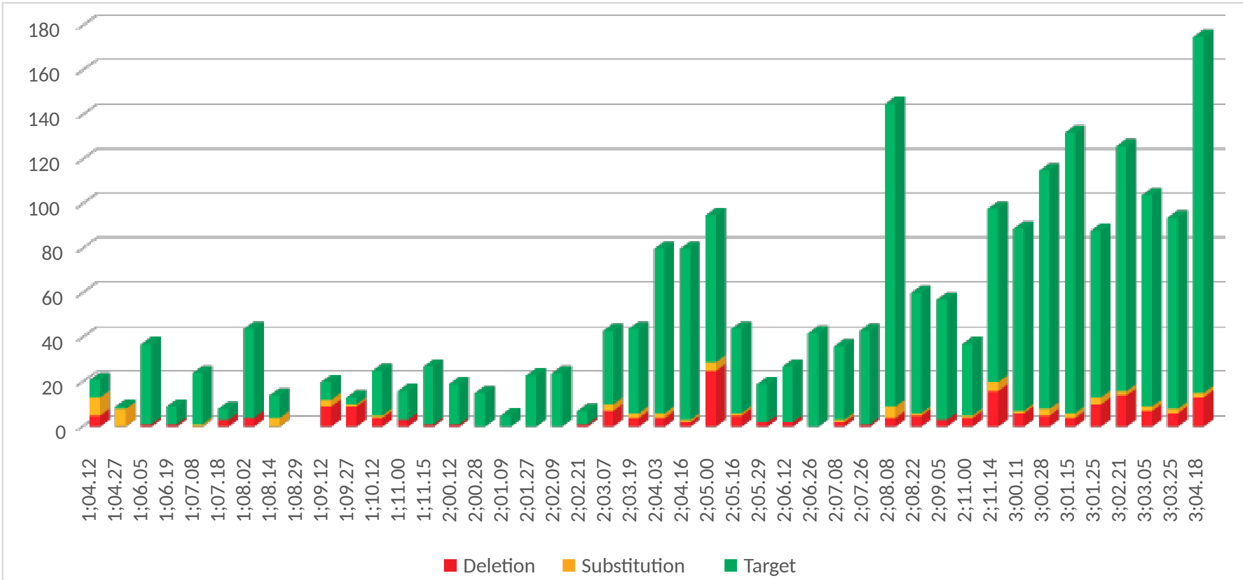


Figure 28: Longitudinal development of word-initial unstressed /w/ in singleton onsets

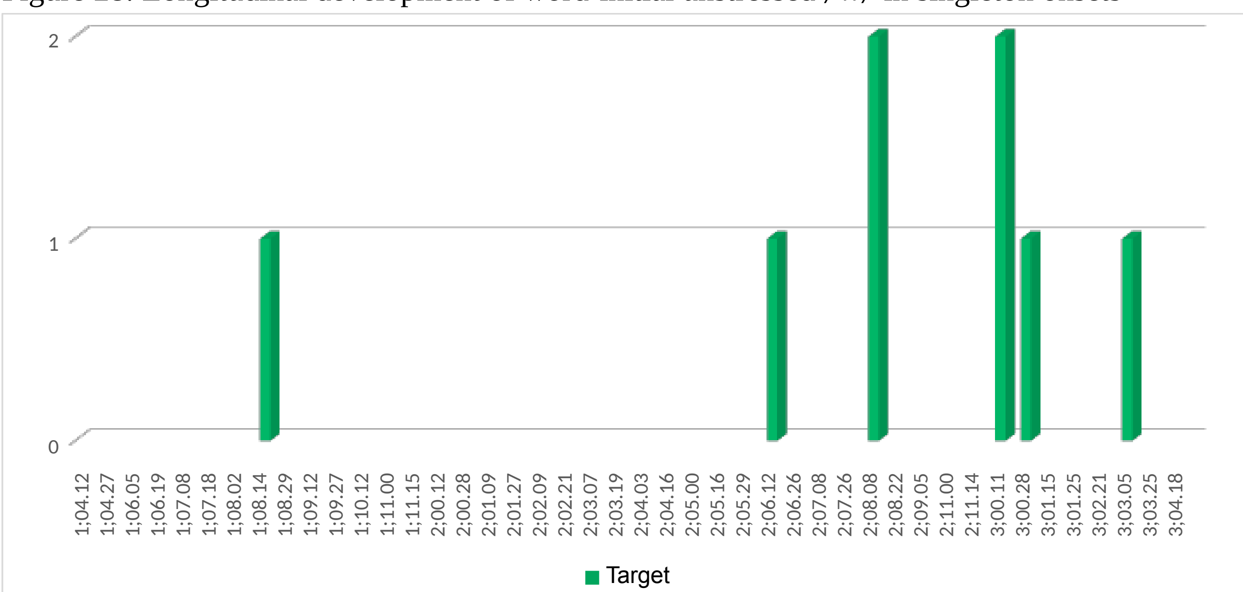


Figure 29: Longitudinal development of word-medial stressed /w/ in singleton onsets

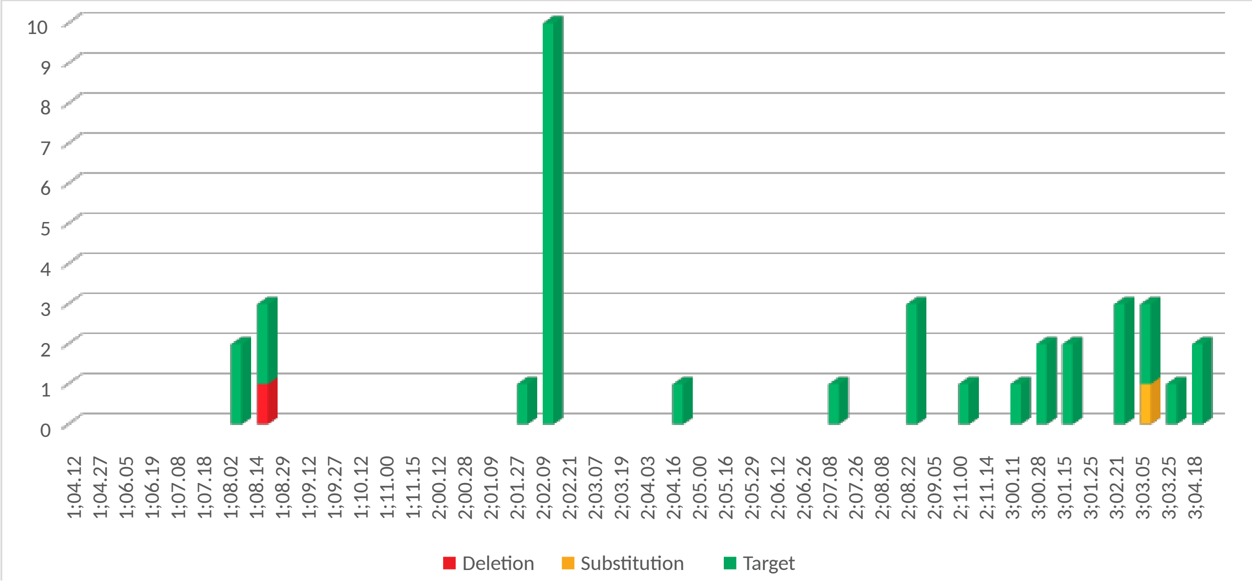


Figure 30: Longitudinal development of word-medial unstressed /w/ in singleton onsets

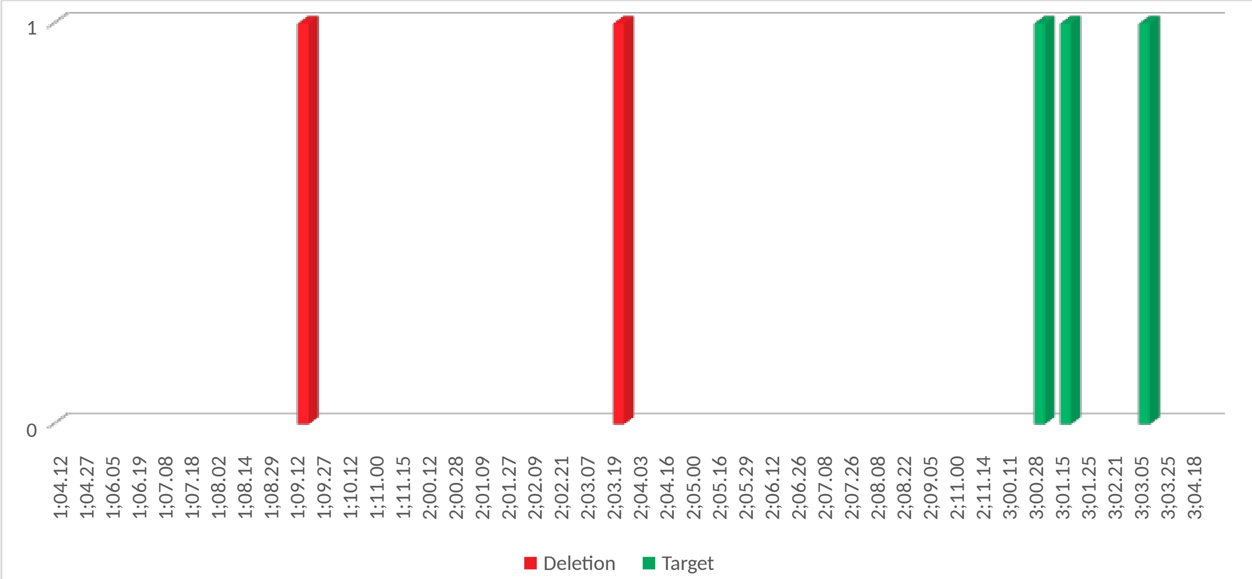


Figure 31: Longitudinal development of word-initial stressed /ɹ/ in singleton onsets

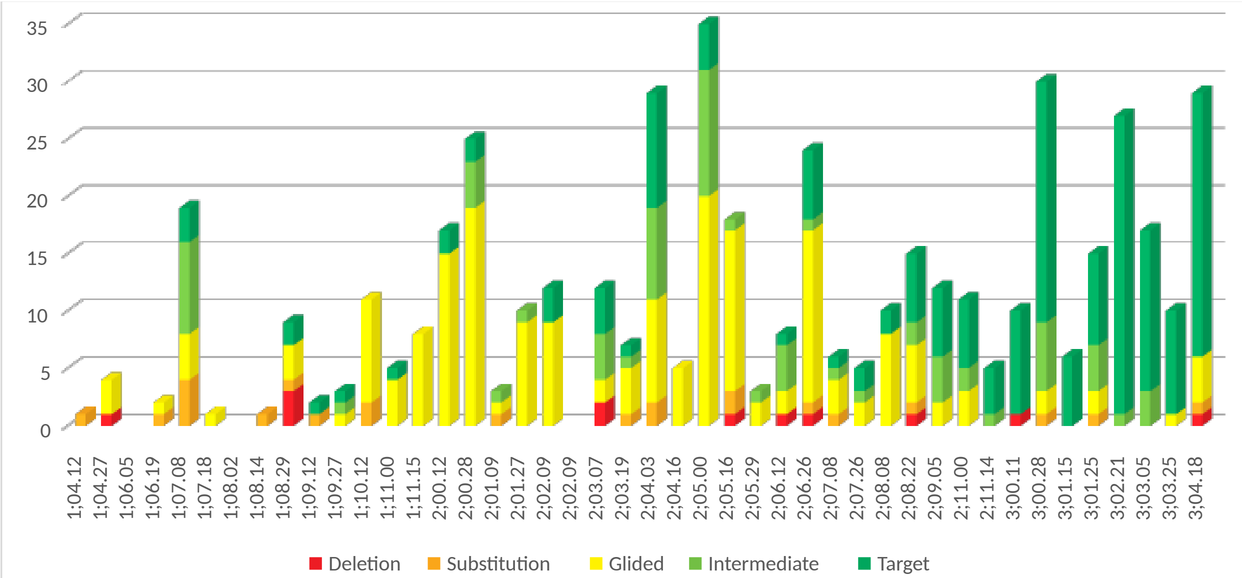


Figure 32: Longitudinal development of word-initial unstressed /ɹ/ in singleton onsets

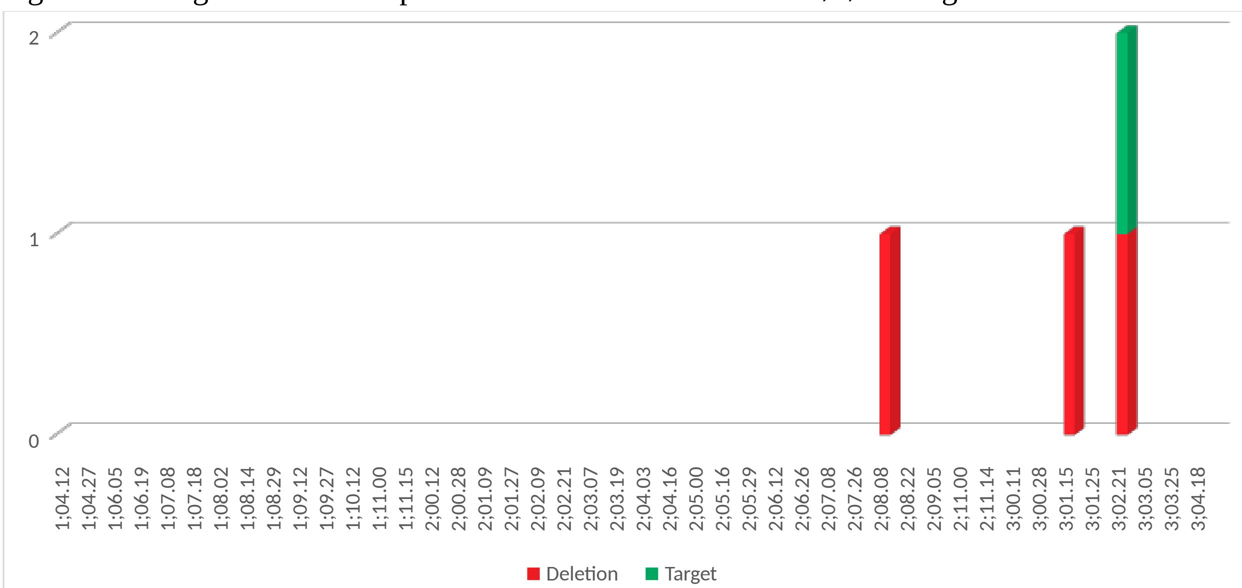


Figure 33: Longitudinal development of word-medial stressed /ɹ/ in singleton onsets

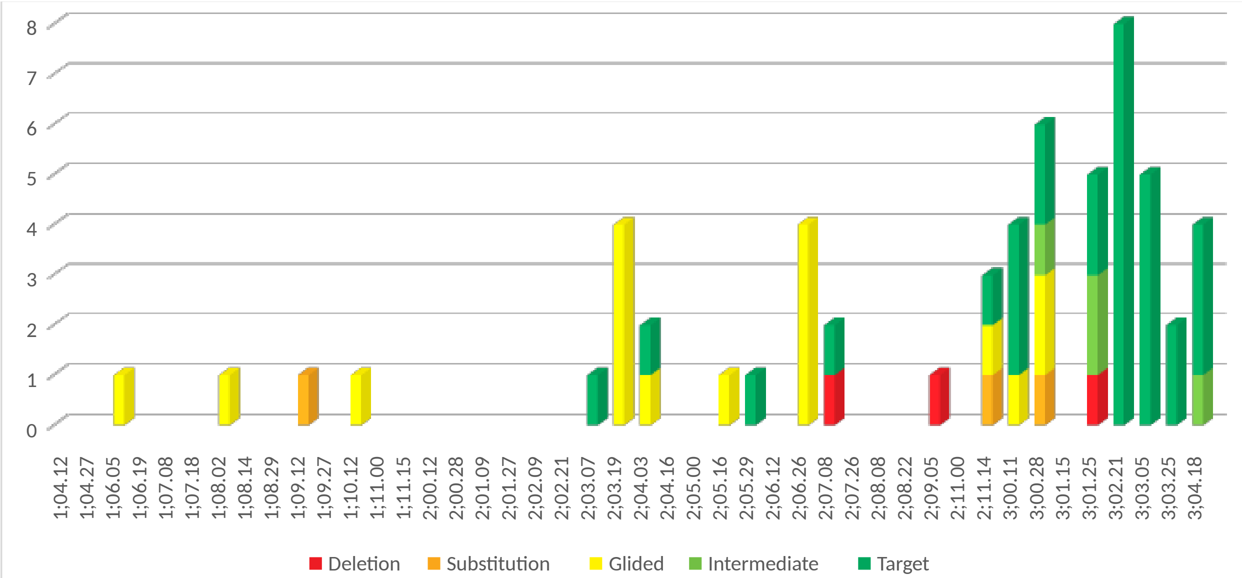


Figure 34: Longitudinal development of word-medial unstressed /ɹ/ in singleton onsets

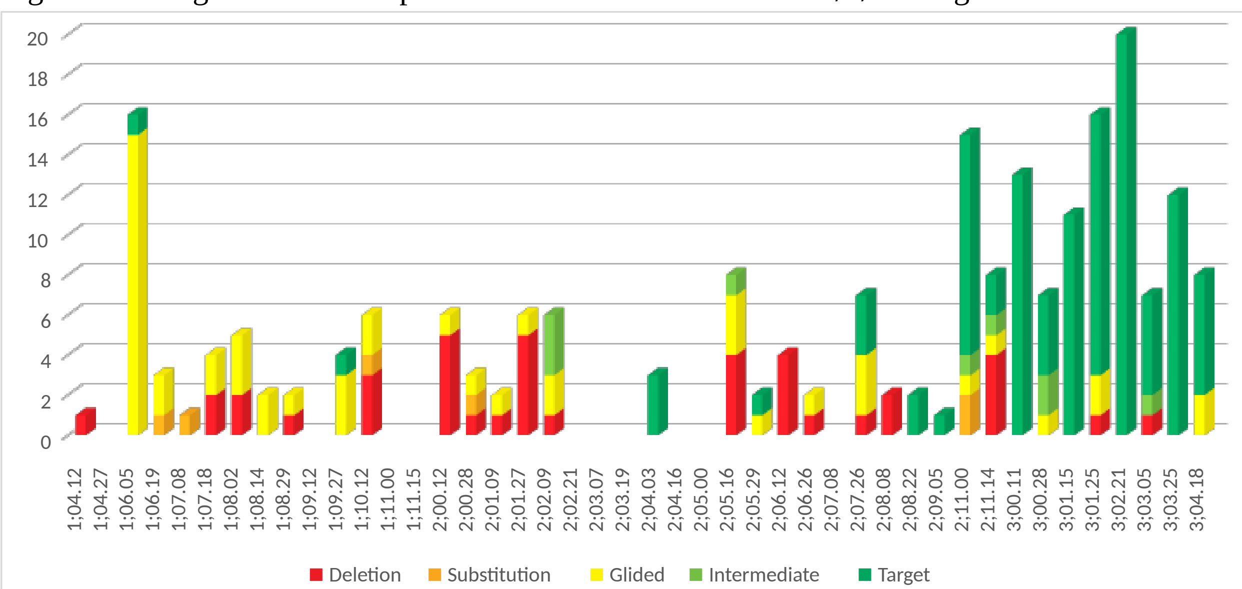


Figure 35: Longitudinal development of word-initial stressed /tw/ onsets

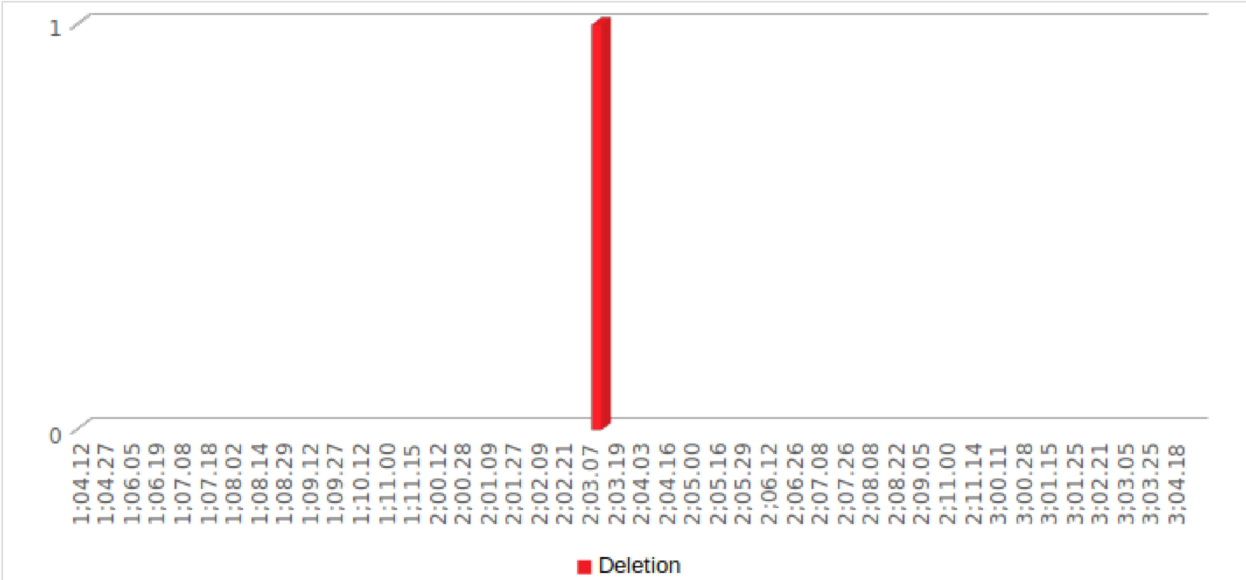


Figure 36: Longitudinal development of word-initial stressed /sw/ onsets

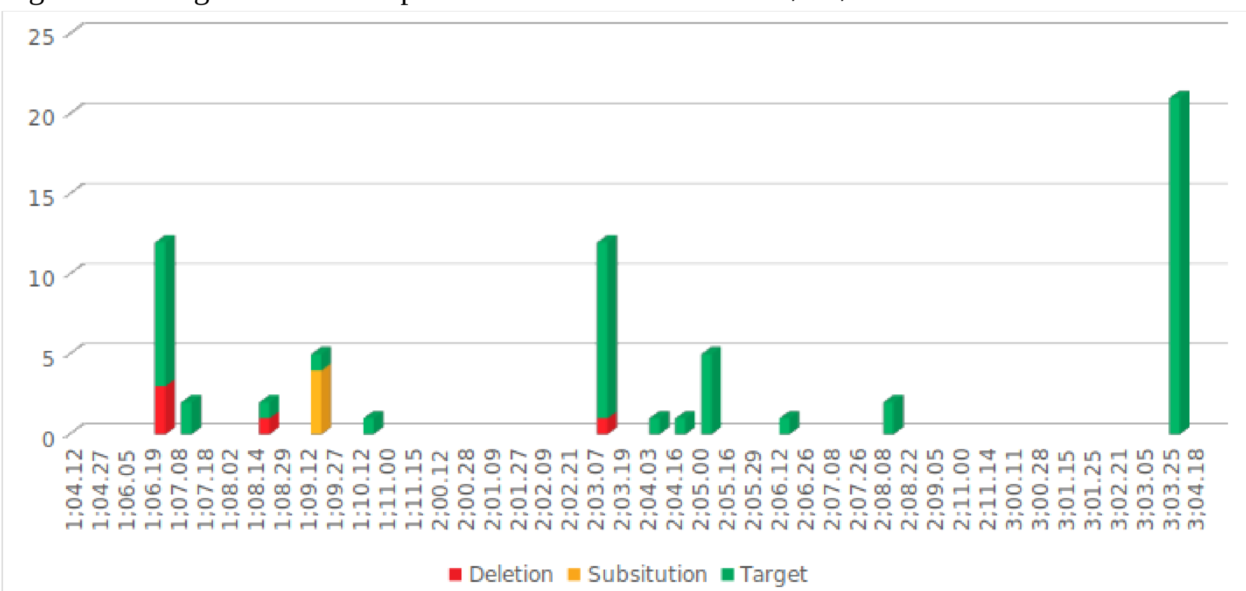


Figure 37: Longitudinal development of word-initial stressed /kw/ onsets

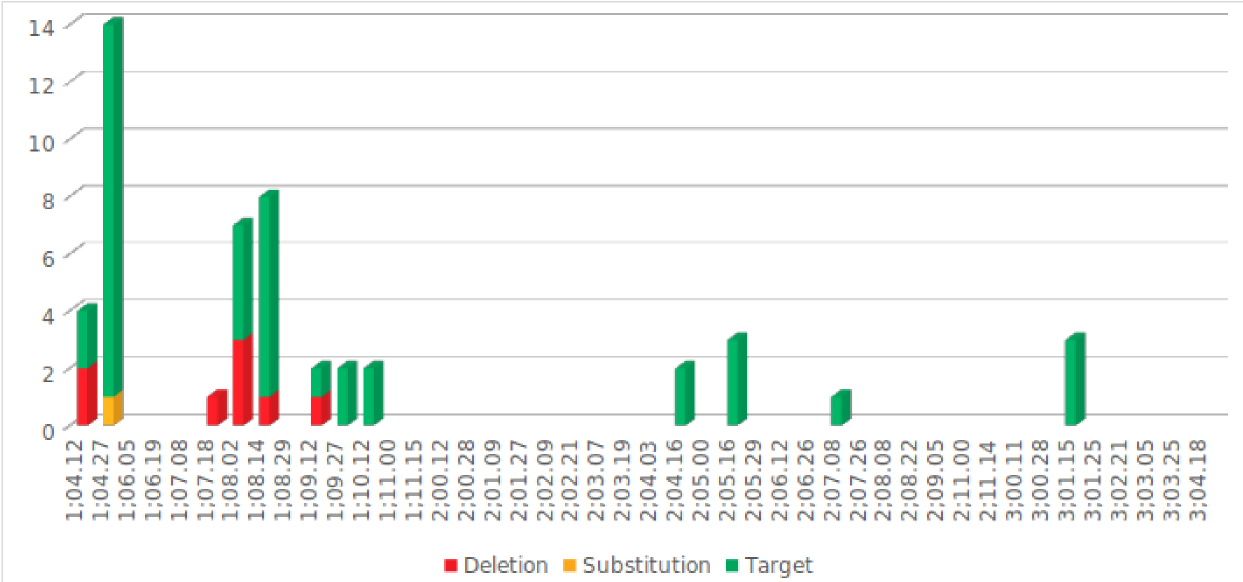


Figure 38: Longitudinal development of word-medial unstressed /kw/ onsets

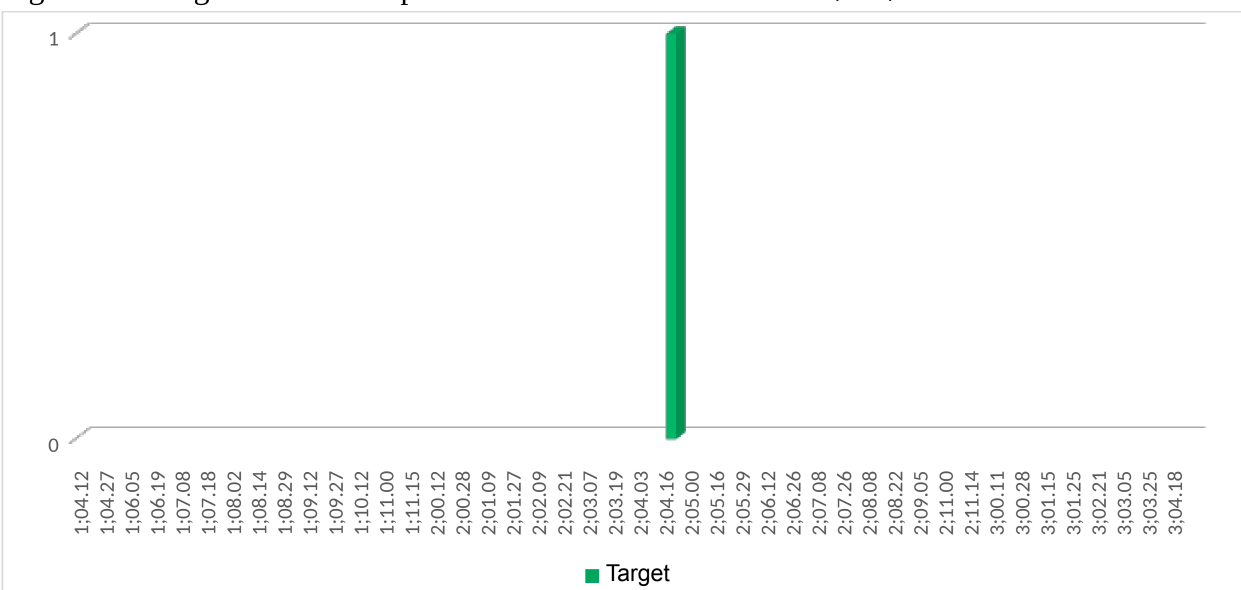


Figure 39: Longitudinal development of word-medial unstressed /gw/ onsets

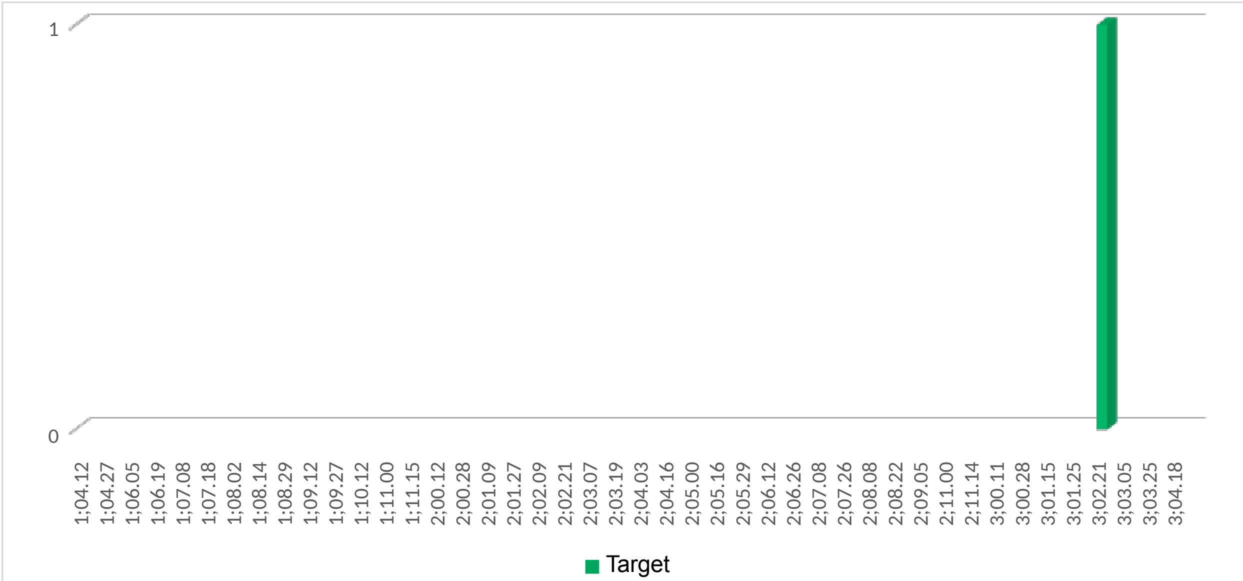


Figure 40: Longitudinal development of word-initial stressed /pɹ/ onsets

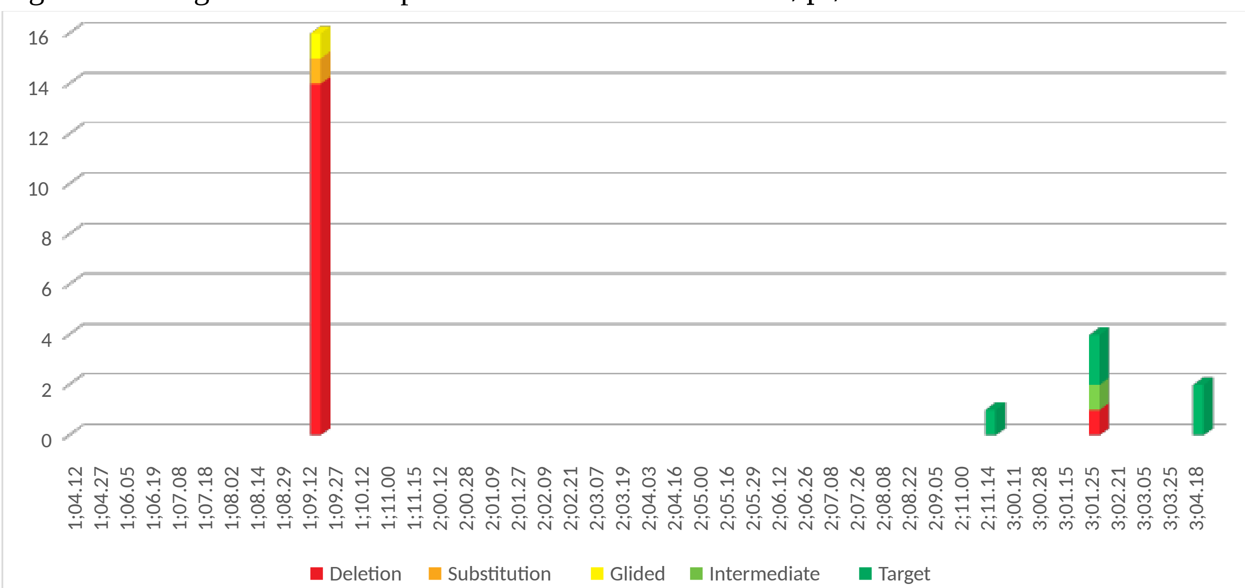


Figure 41: Longitudinal development of word-initial stressed /bɹ/ onsets

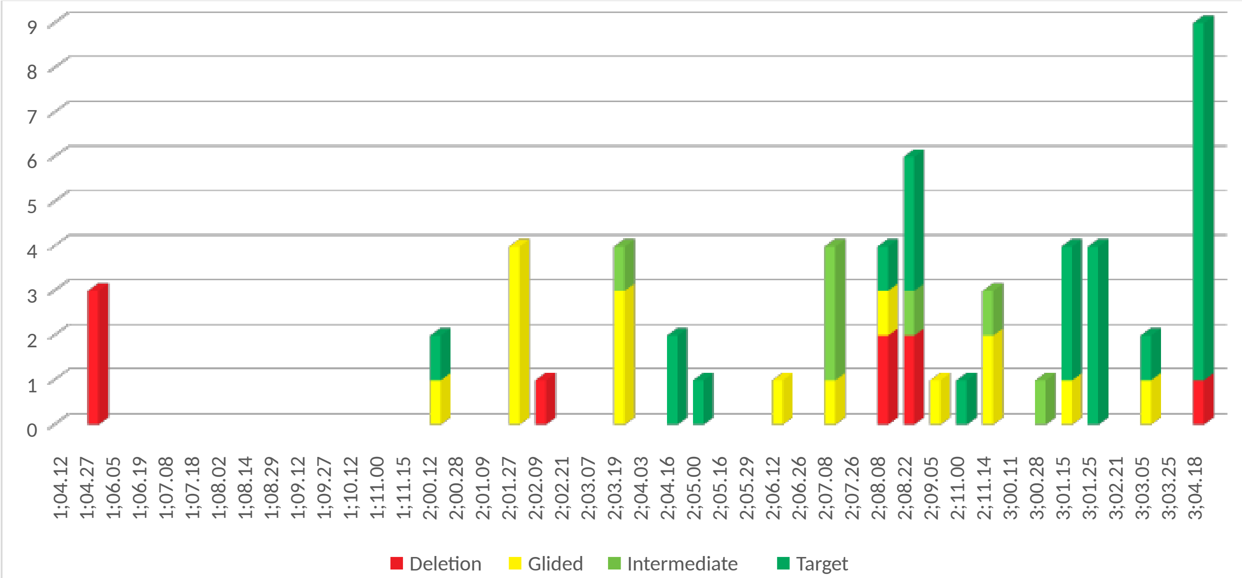


Figure 42: Longitudinal development of word-medial stressed /bɹ/ onsets

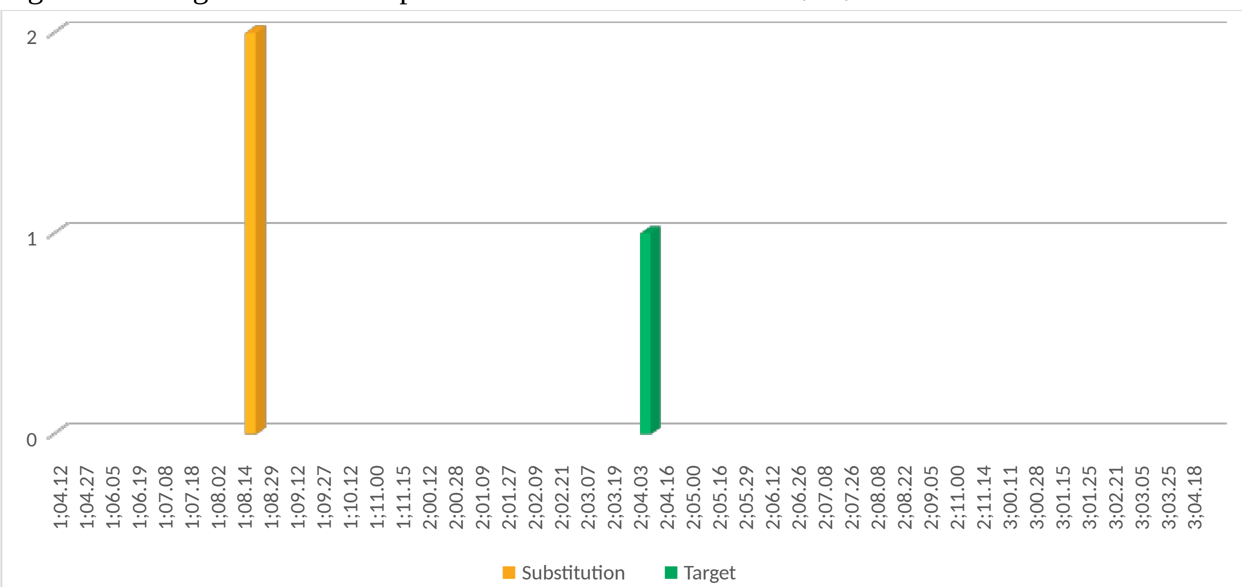


Figure 43: Longitudinal development of word-medial unstressed /bɹ/ onsets

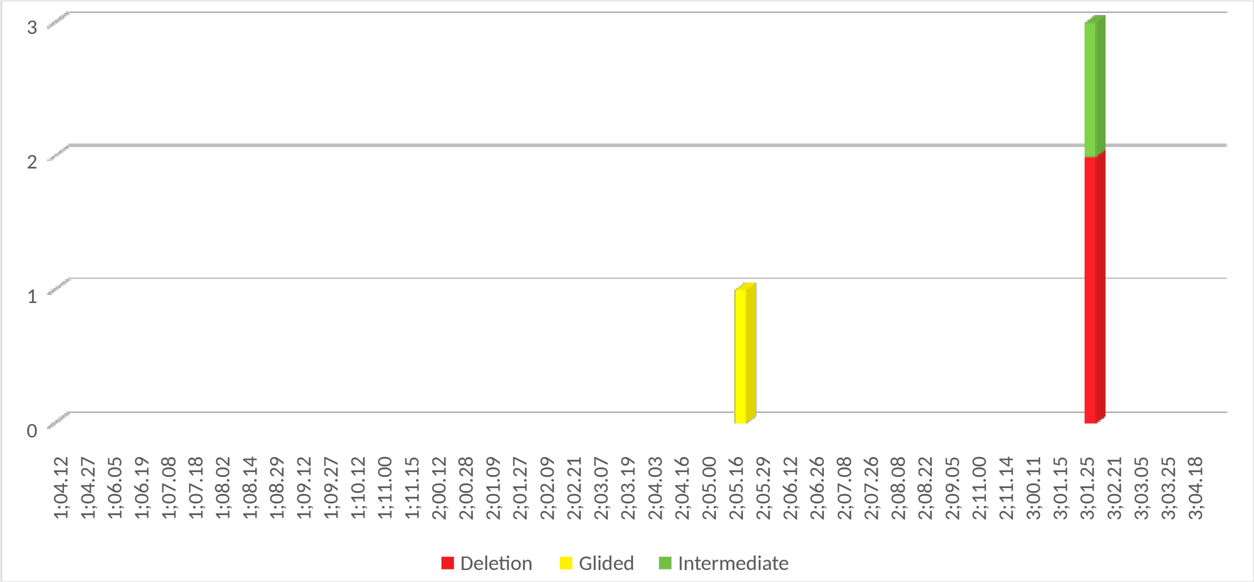


Figure 44: Longitudinal development of word-initial stressed /tɹ/ onsets

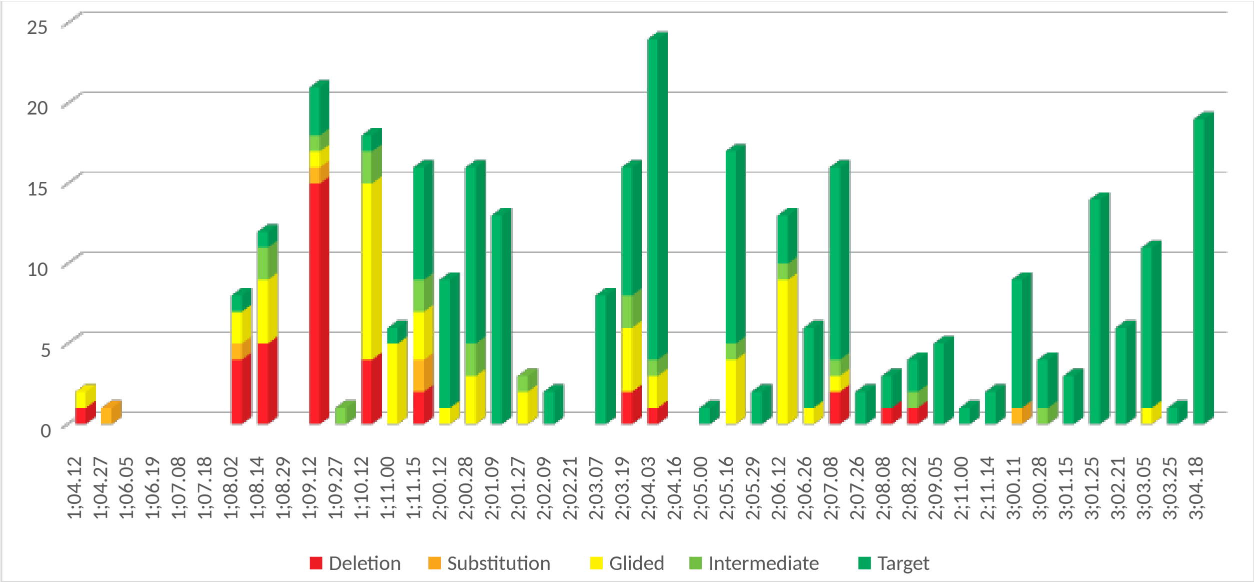


Figure 45: Longitudinal development of word-medial stressed /tɹ/ onsets

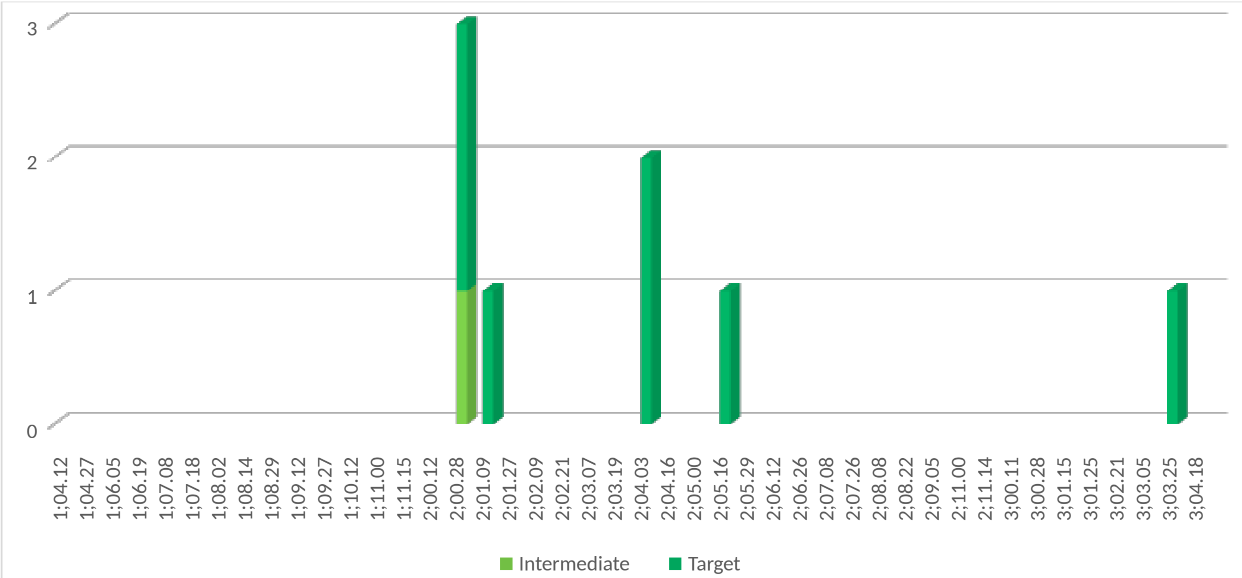


Figure 46: Longitudinal development of word-medial unstressed /tɹ/ onsets

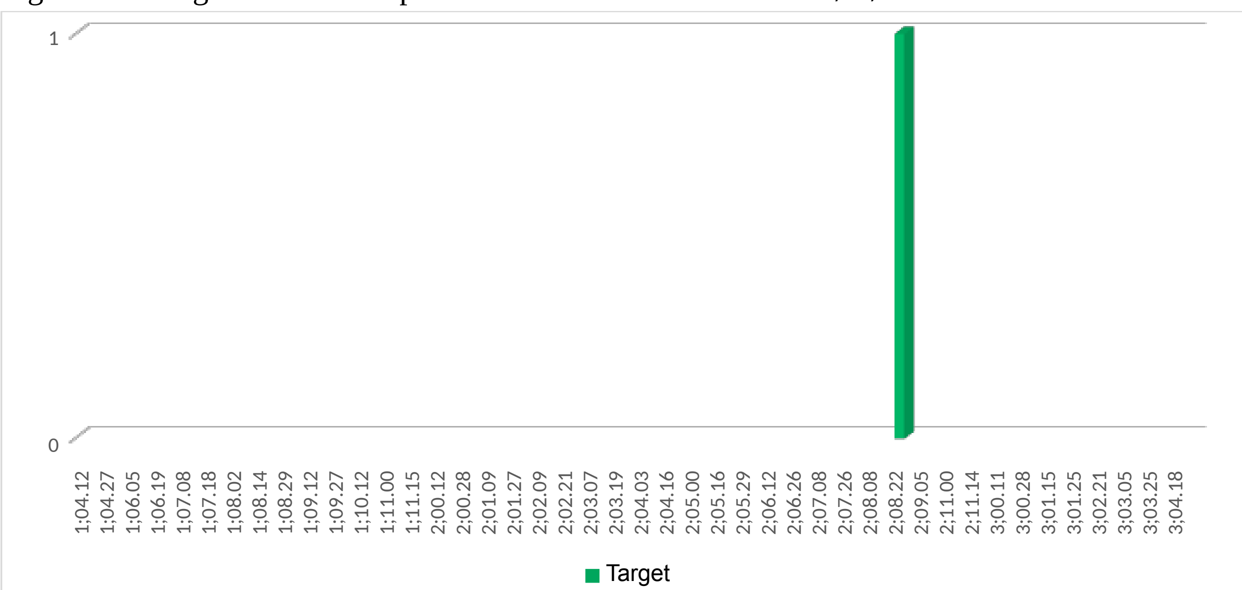


Figure 47: Longitudinal development of word-initial stressed /dɹ/ onsets

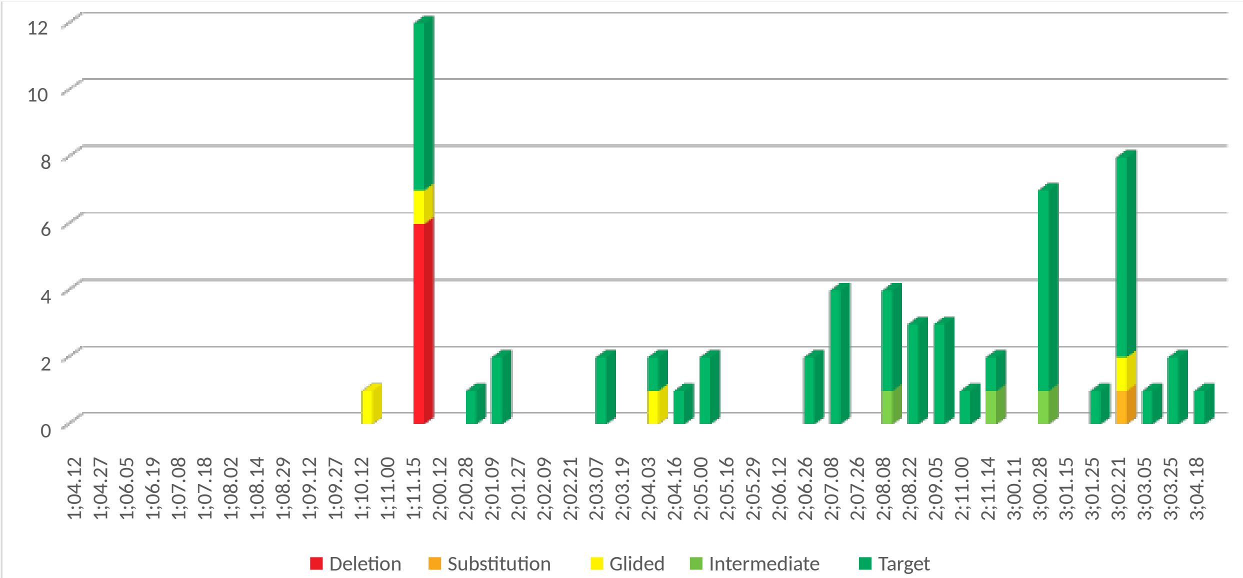


Figure 48: Longitudinal development of word-medial unstressed /dɹ/ onsets

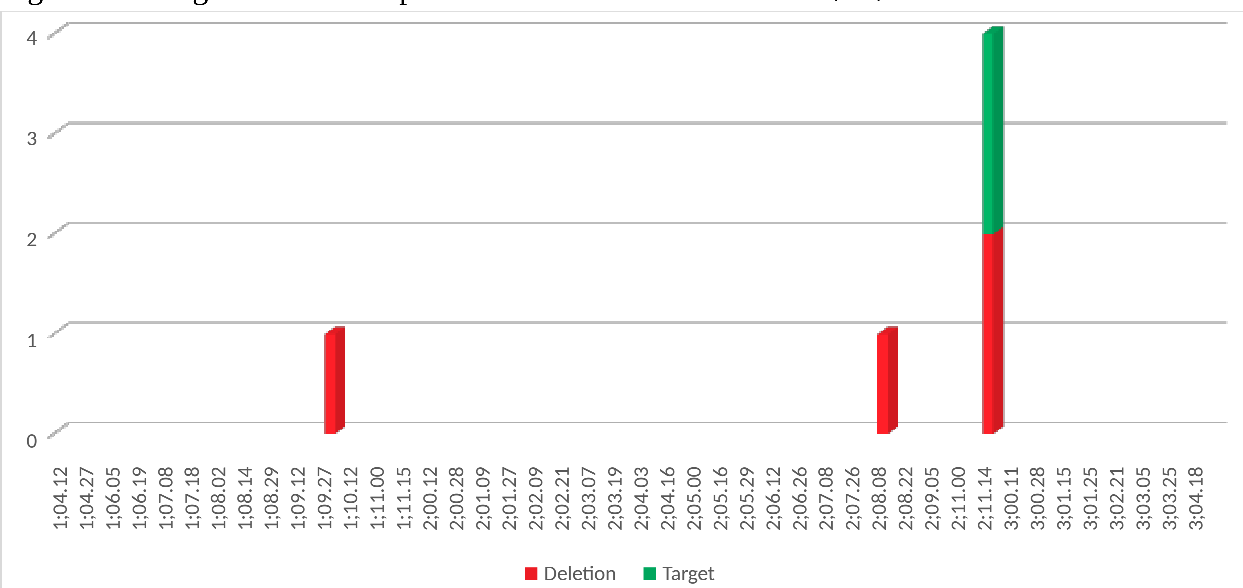


Figure 49: Longitudinal development of word-initial stressed /kɹ/ onsets

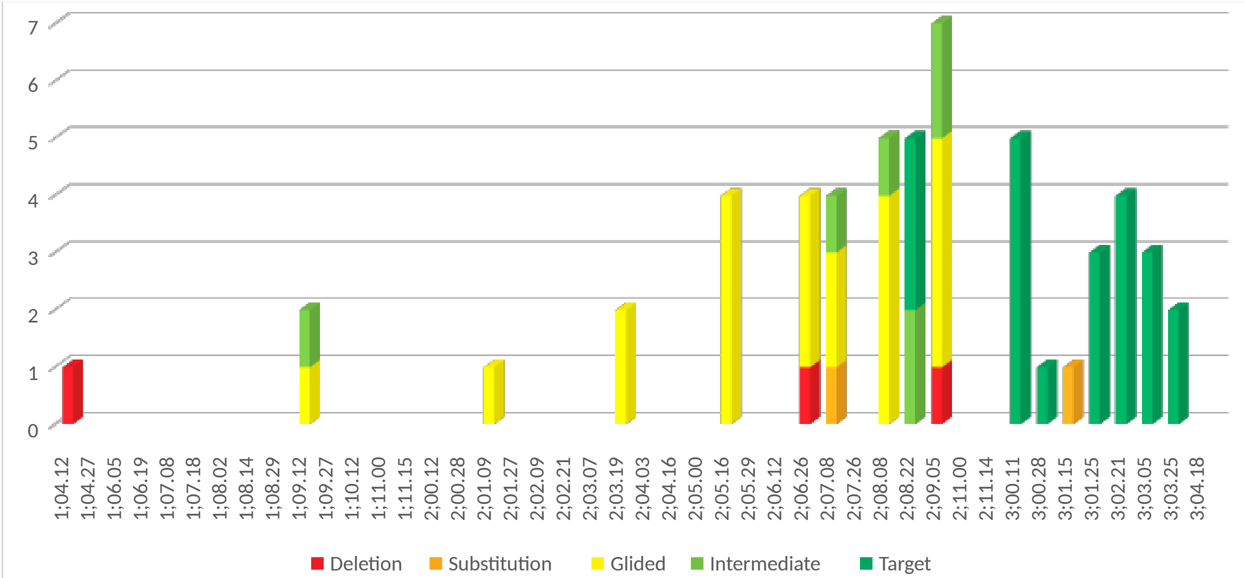


Figure 50: Longitudinal development of word-medial stressed /kɹ/ onsets

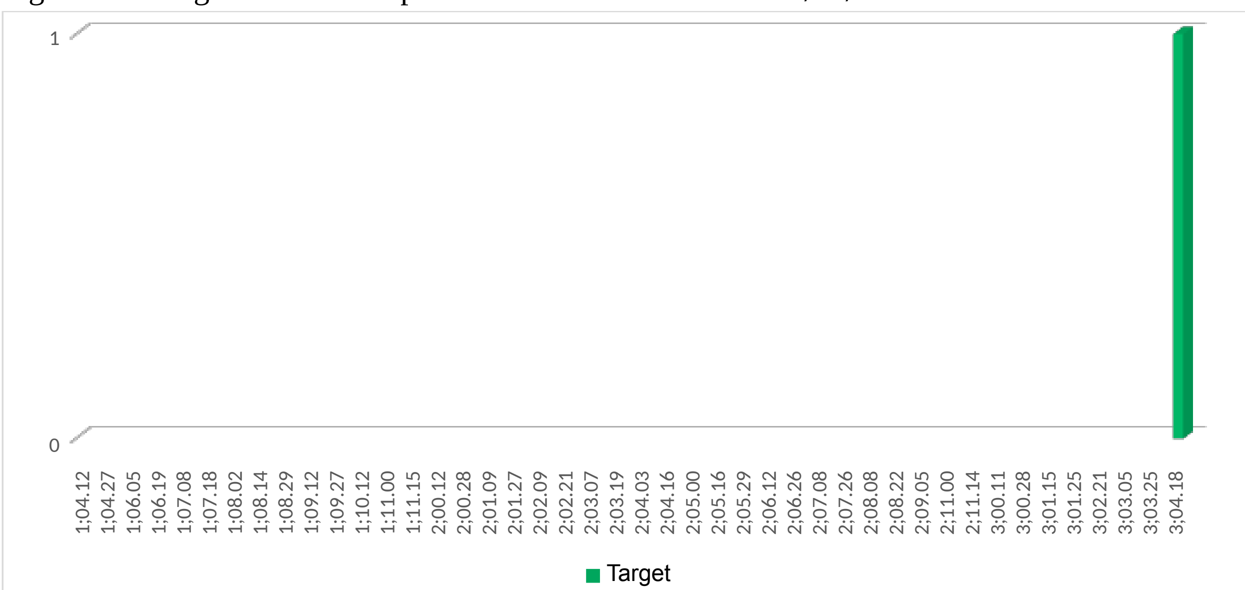


Figure 51: Longitudinal development of word-medial unstressed /kɹ/ onsets

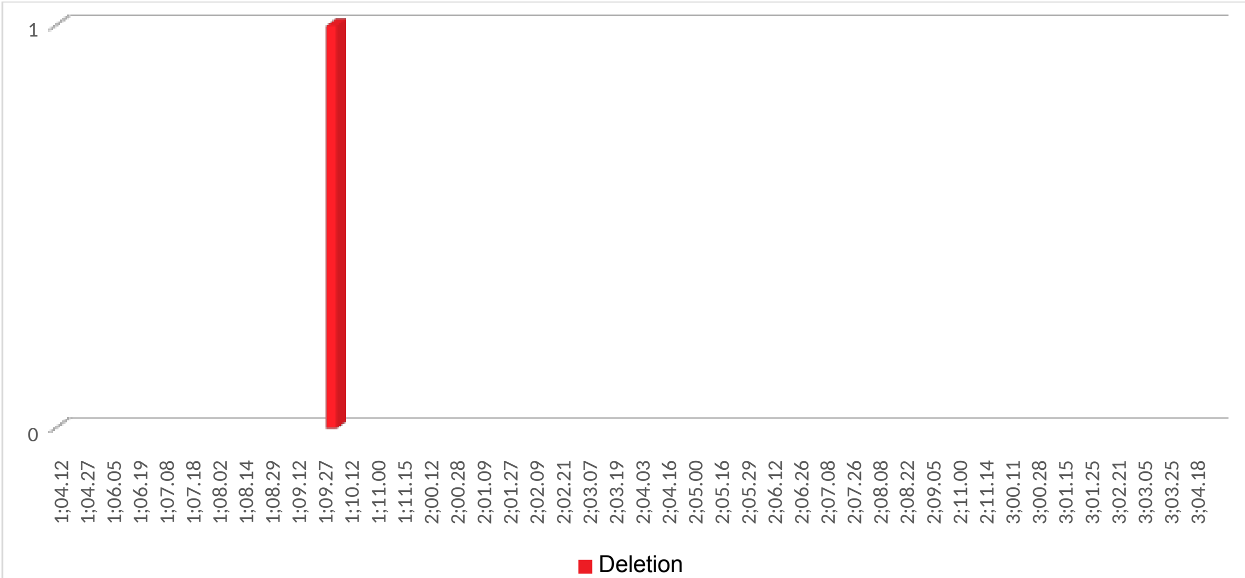


Figure 52: Longitudinal development of word-initial stressed /gɹ/ onsets

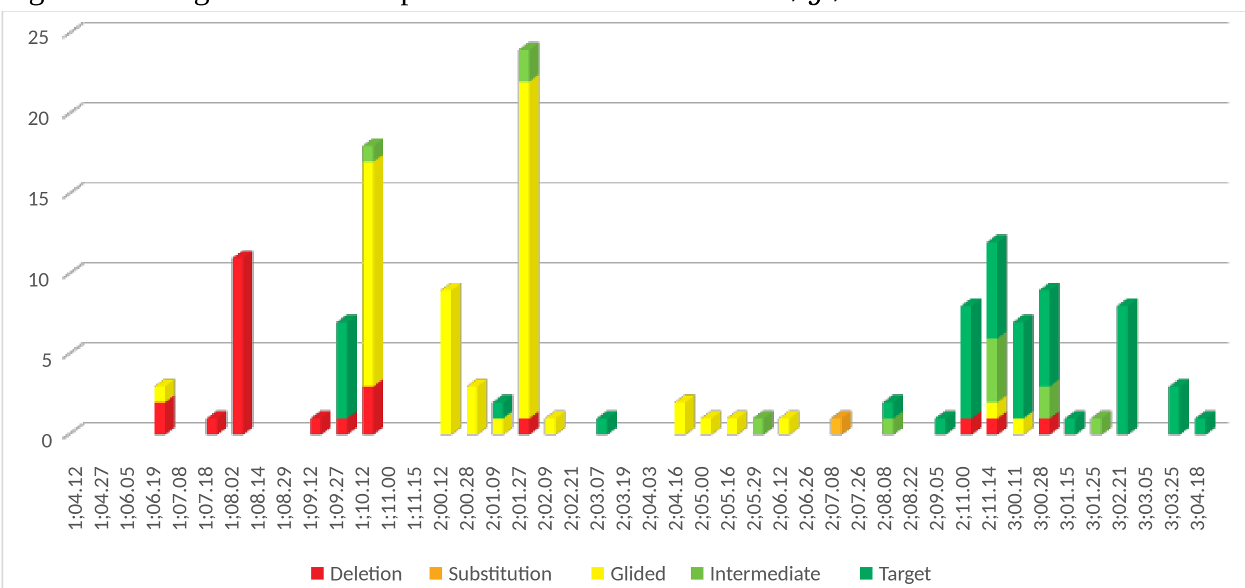


Figure 53: Longitudinal development of word-medial stressed /gɹ/ onsets

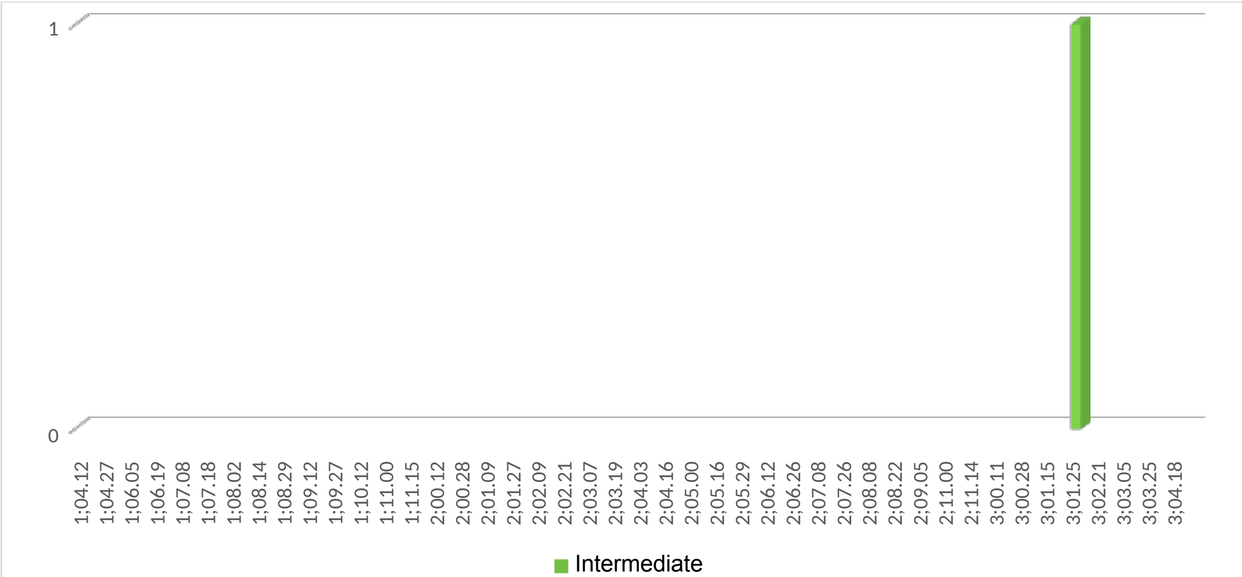


Figure 54: Longitudinal development of word-medial unstressed /gɹ/ onsets

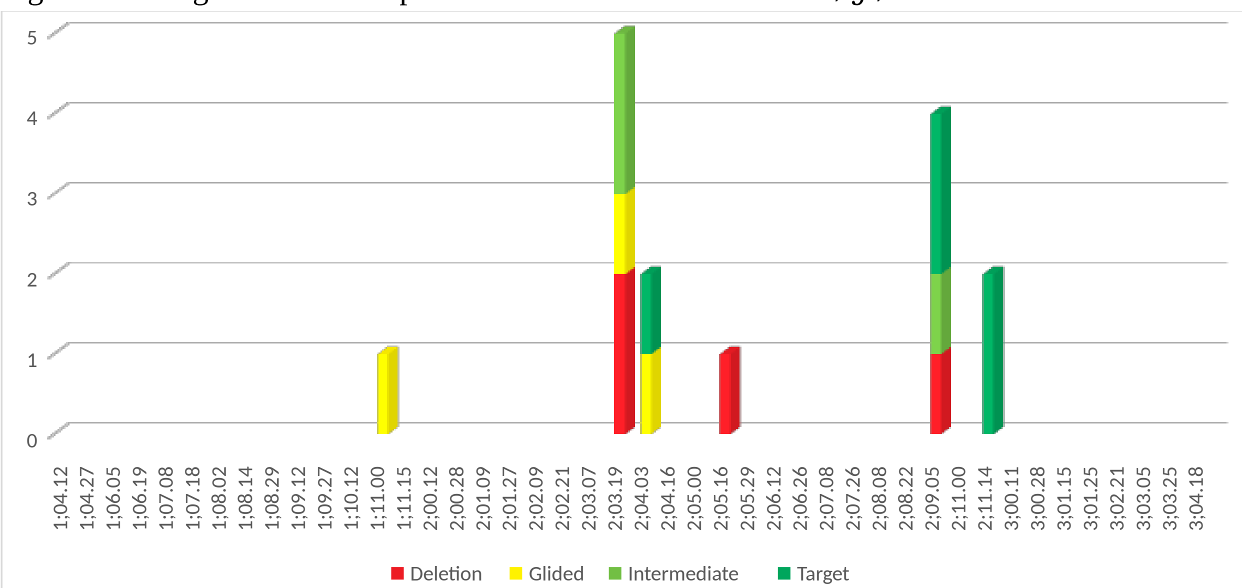


Figure 55: Longitudinal development of word-initial stressed /fɹ/ onsets

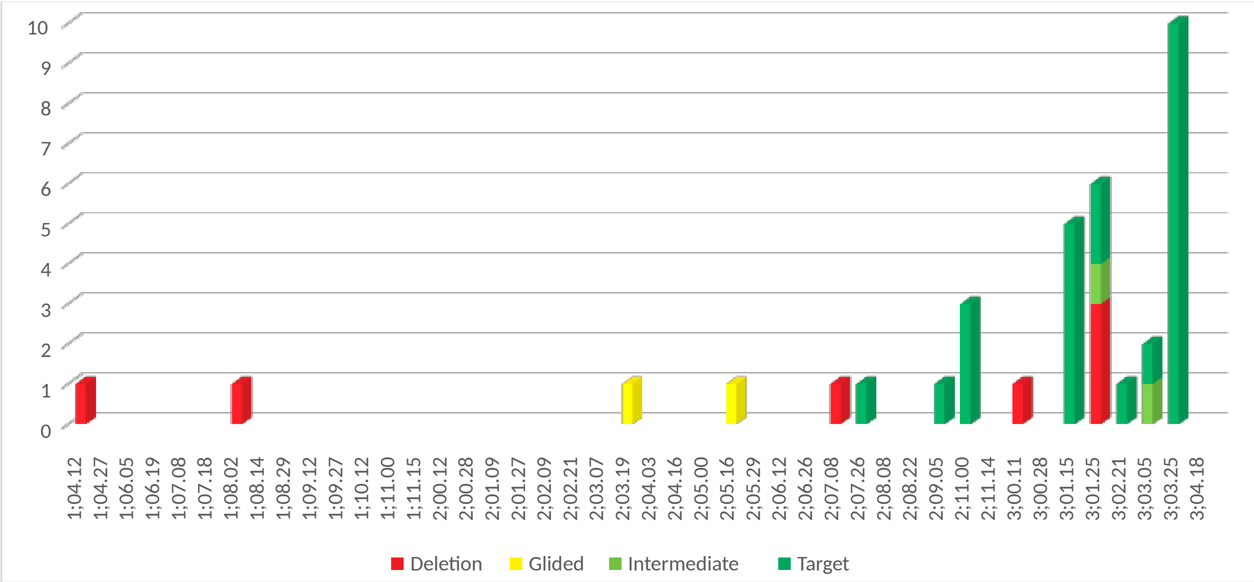


Figure 56: Longitudinal development of word-initial stressed /vɹ/ onsets

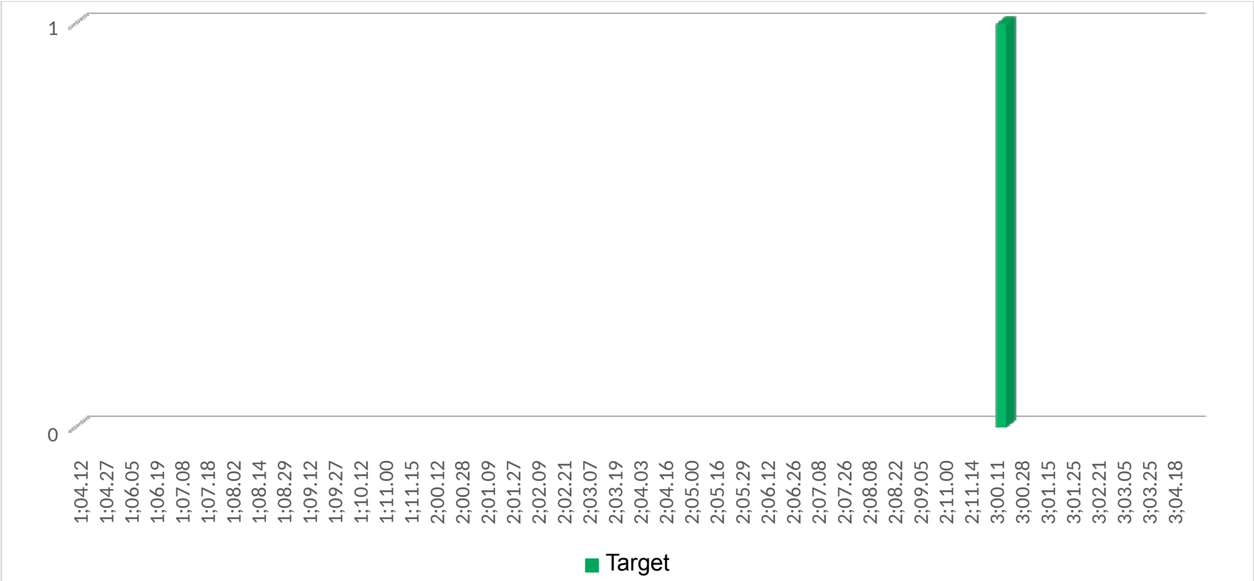


Figure 57: Longitudinal development of word-medial unstressed /vɹ/ onsets

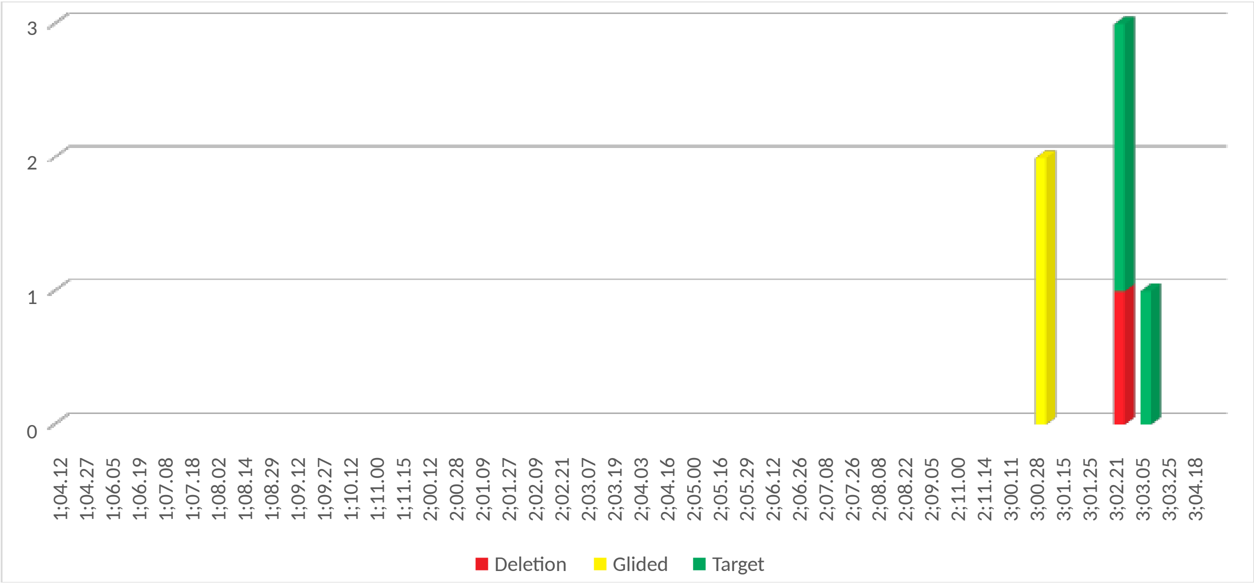
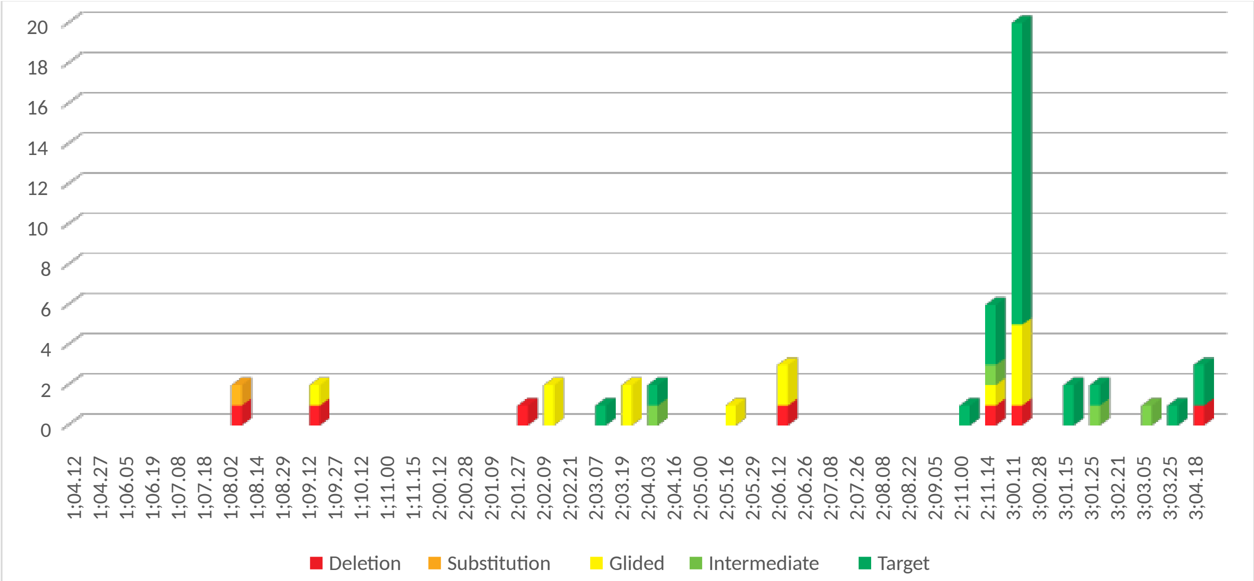


Figure 58: Longitudinal development of word-initial stressed /θɹ/ onsets



Appendix B: Standard deviation charts

Figure 59: Standard deviation of formants for singleton /ɹ/ perceived as [ɹ]

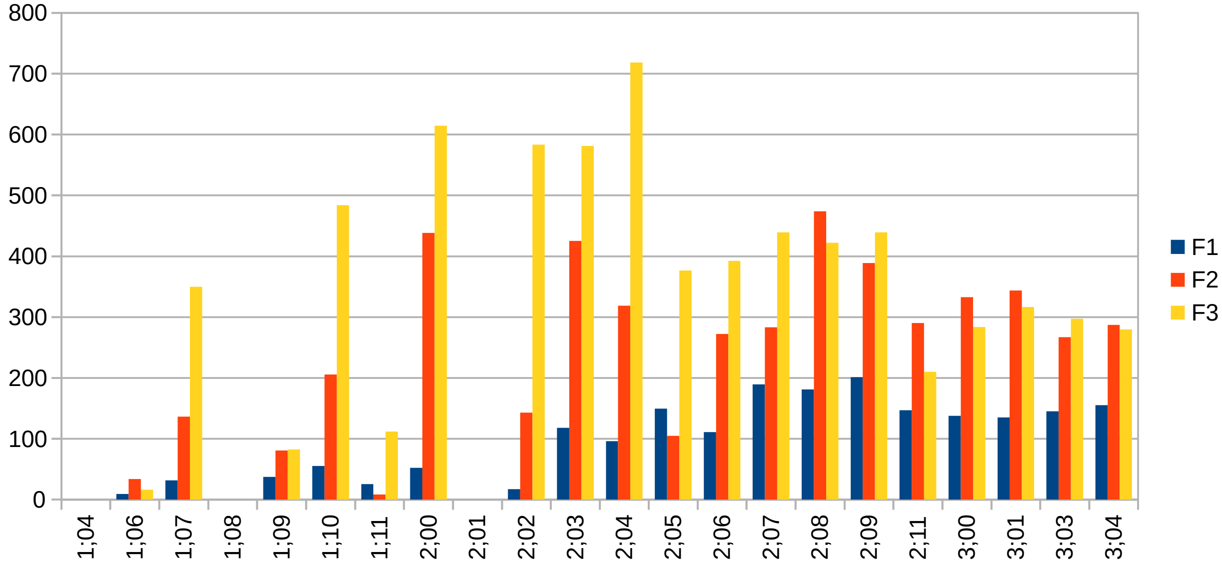


Figure 60: Standard deviation of formants for singleton /ɹ/ perceived as [w]

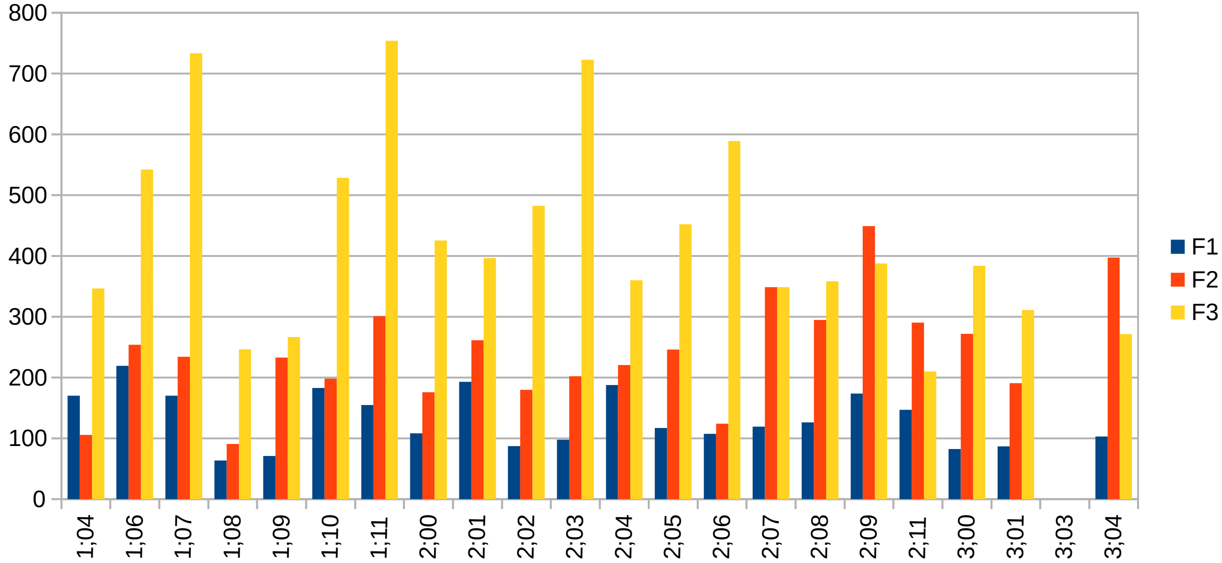


Figure 61: Standard deviation of formants for singleton /ɹ/ perceived as an intermediate pronunciation

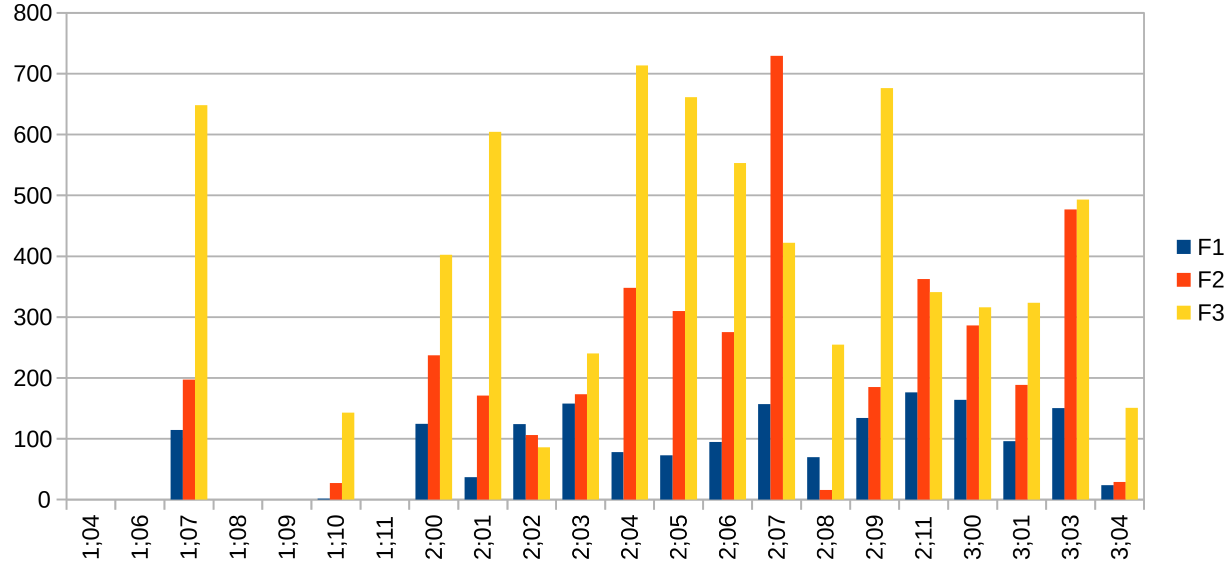


Figure 62: Standard deviation of formants for singleton [w]

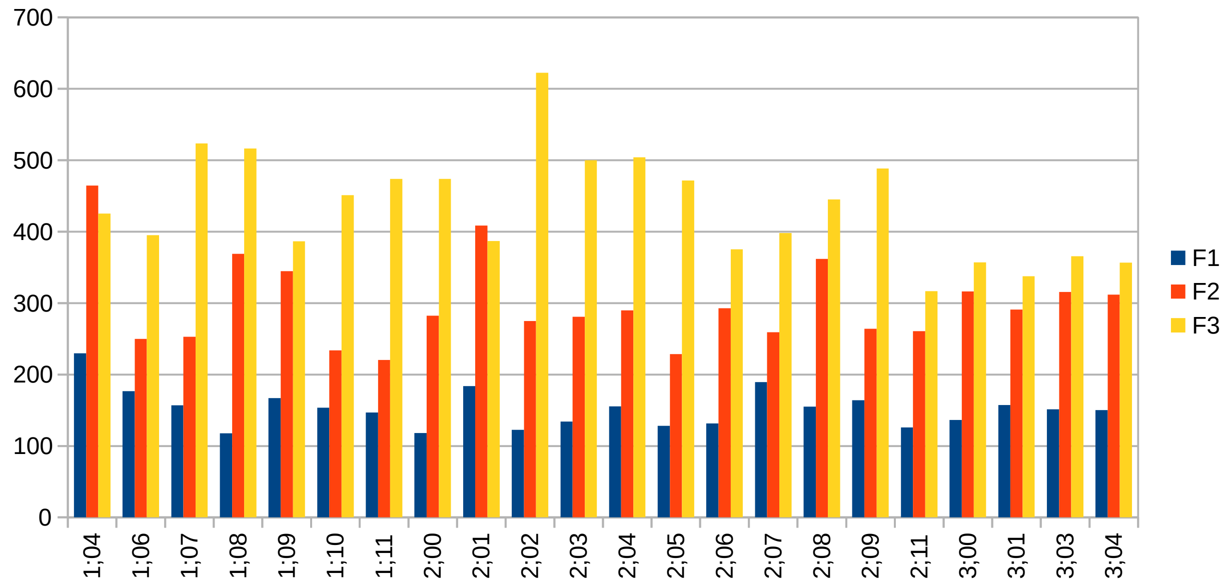


Figure 63: Standard deviation of formants for rhotic /ɹ/ in labial-initial complex onsets

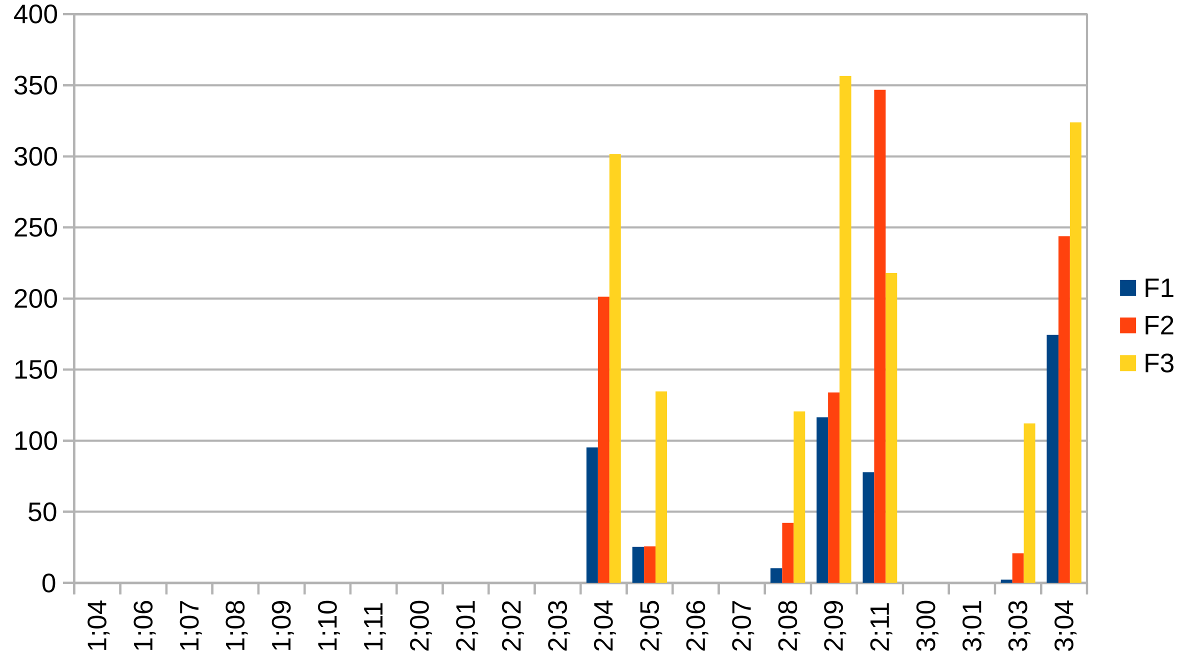


Figure 64: Standard deviation of formants for labialized /ɹ/ in labial-initial complex onsets

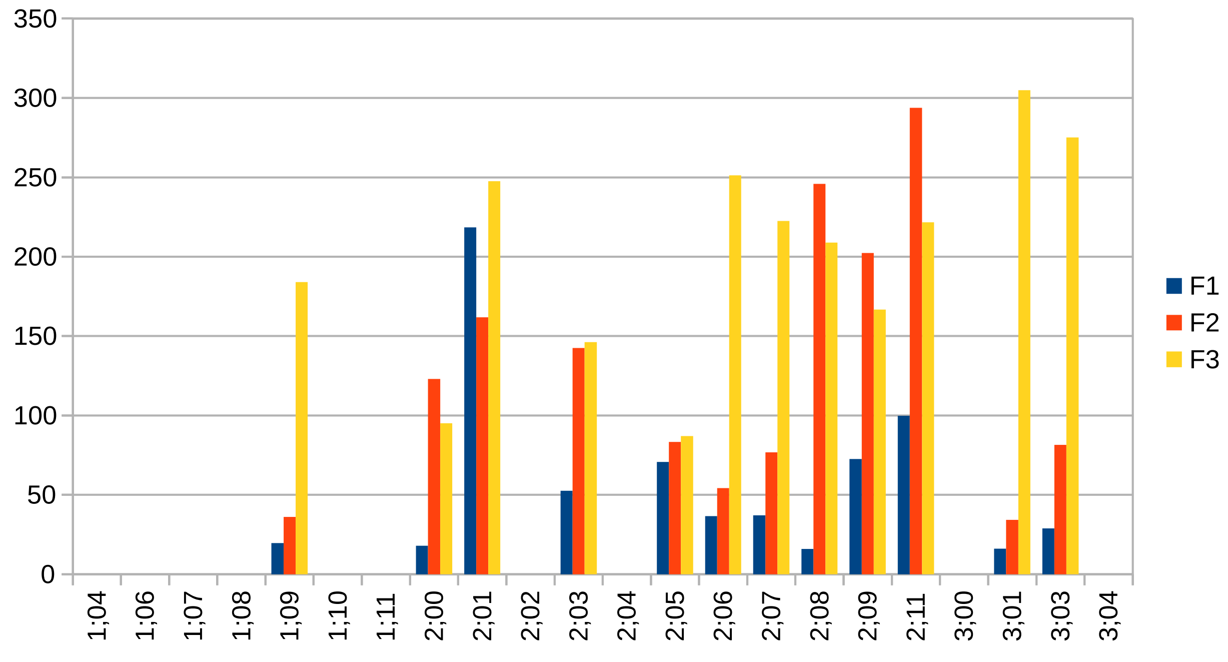


Figure 65: Standard deviation of formants for rhotic /ɹ/ in coronal stop-initial complex onsets

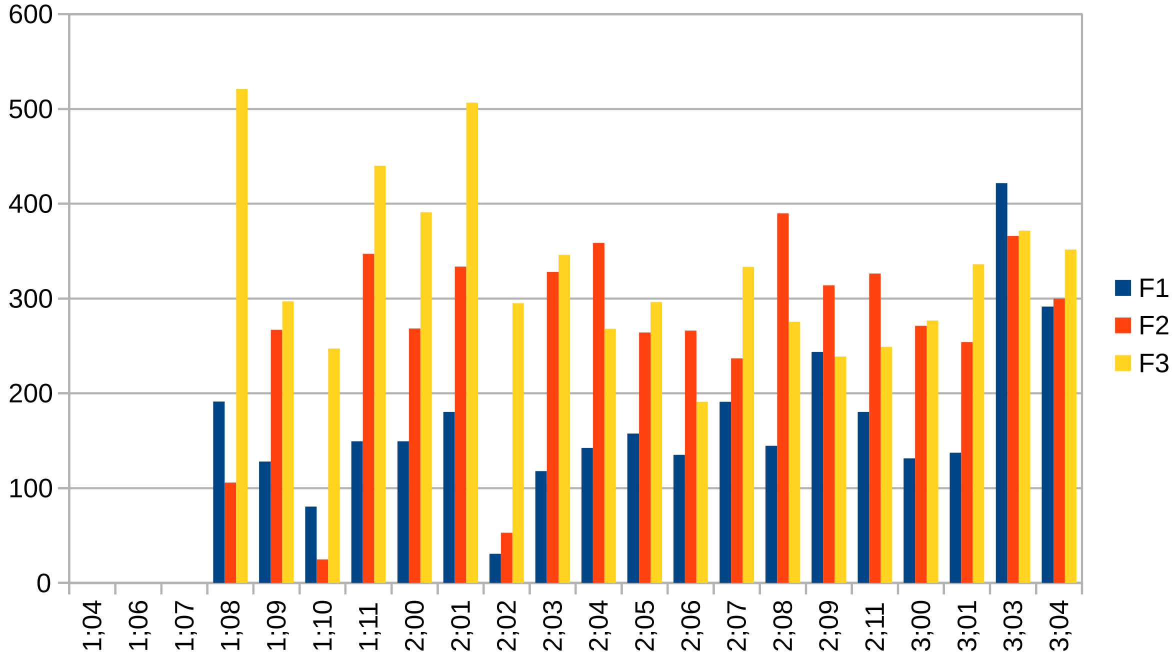


Figure 66: Standard deviation of formants for labialized /ɹ/ in coronal stop-initial complex onsets

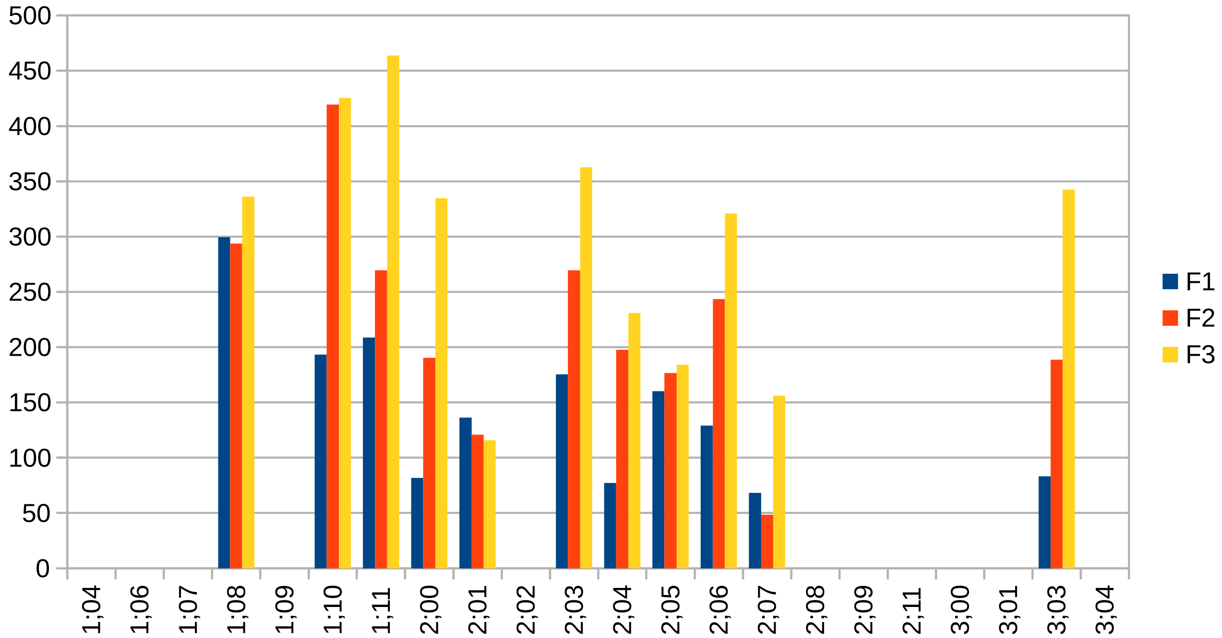


Figure 67: Standard deviation of formants for /ɹ/ in velar-initial complex onsets

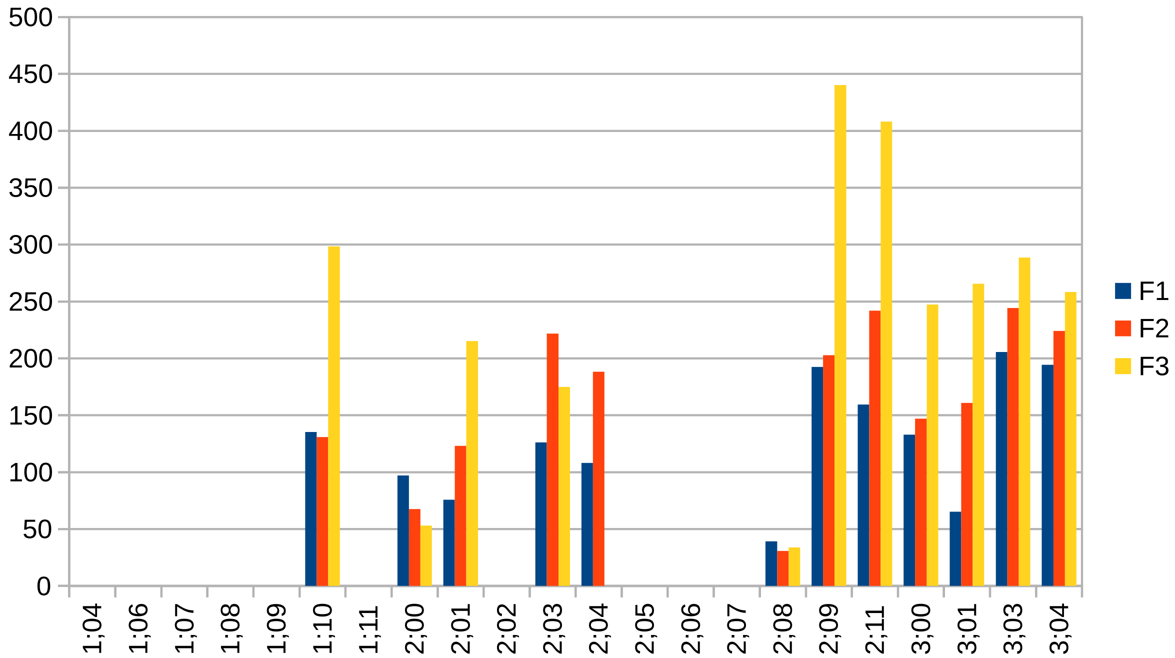


Figure 68: Standard deviation of formants for /w/ in coronal-initial complex onsets

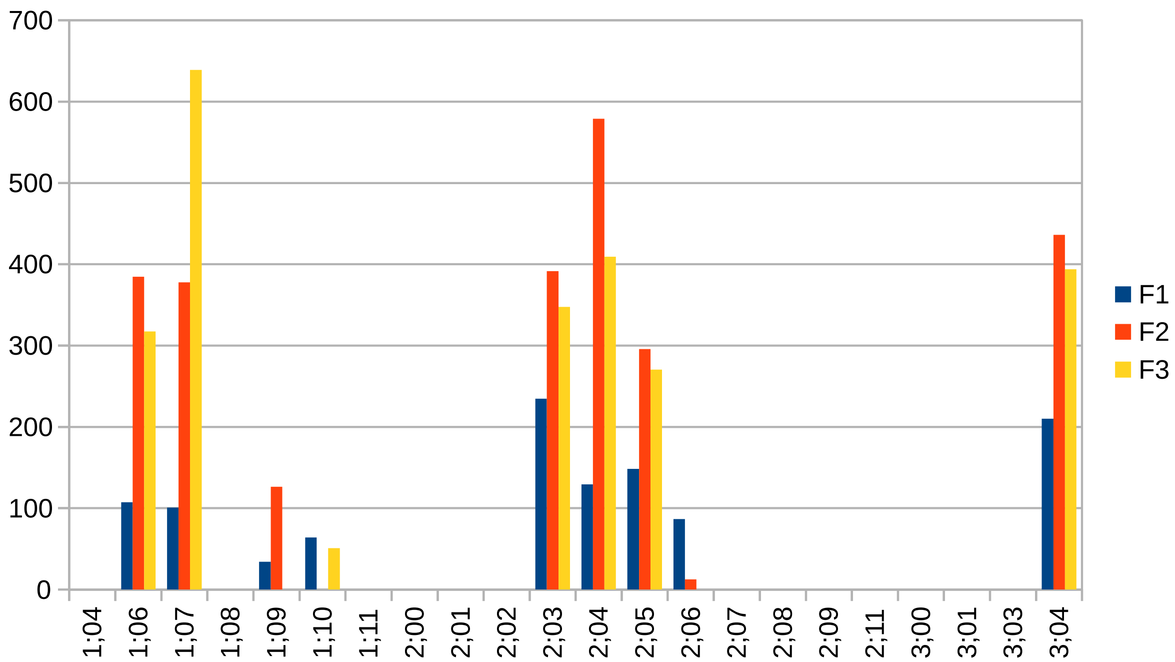


Figure 69: Standard deviation of formants for /w/ in velar-initial complex onsets