SPEECH PERCEPTION IN CHILDREN WITH READING DISABILITIES

BY

JIE LU

DISSERTATION

Submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in Nuclear, Plasma, and Radiological Engineering
in the Graduate College of the
University of Illinois at Urbana-Champaign, 2018

Urbana, Illinois

Doctoral Committee:

Professor Jont Allen, Chair
Professor Rizwan Uddin
Associate Professor Cynthia J. Johnson
Associate Professor Robert Wickesberg
Professor Elizabeth Stine-Morrow
Assistant Professor Yang Zhang

ABSTRACT

Reading disability (RD) is typically viewed as a major obstacle in the development of literacy. This thesis proposes that the source of RDs is related to inadequate phonetic non-categorical processing skills rooted in early development. This view is supported by two experiments on children with reading disabilities in two tasks: a Syllable Confusion Oddball task (SCO) and a Nonsense Syllable Confusion Matrix task (NSCM). The SCO task tested children's ability to select a different syllable (either a consonant vowel–CV or vowel consonant–VC syllable) from a string of three such syllables spoken by three different talkers. The NSCM task tested children's ability to reproduce the exact syllable after hearing it. All of the natural CV and VC speech syllables used in both tasks were taken from a commercial database of 18 talkers.

Experimental results were: First, regardless of pure tone hearing ability and higher language processing ability (at the level of words and sentences), children with RD (RDs) encountered significant difficulties in phonetic perception compared to a normal reading control group (RCs). Second, RDs had a speech perception problem with nonsense syllable identification, despite normal hearing for pure tones. The stark contrast in performance of the two tasks showed that the SCO task was more difficult than the NSCM task for the RDs. Based on this contrast, the hypothesis is that RDs are not able to retain the three syllables in their phonetic short-term memory, as required to produce an accurate outcome. Third, the clustering analysis for probability of error and entropy showed that the RDs had much greater diversity in responses than the RCs. Some RDs gave all possible responses of phones (maximum entropy), while a few RDs gave only one alternative guess for certain phones. It is reasonable to conclude that, while weakness in phonetic perception could separate the majority of the RDs from the RCs, there exist other factors which contribute to the difficulties in reading for some RDs. Finally, major similarities in confusion patterns for the two groups were plotted using directed graphs and stacked bar plots. For consonants, centers of confusion were mainly affricates and fricatives. The RDs also had additional confusion patterns related to stops and liquids. For vowels, perception of diphthongs and tense vowels was mostly intact for both groups. Confusions mostly occurred among certain "front" and "back" lax vowels. With the purpose of developing an automatic diagnostic tool with confusion pattern, methods were explored to rank phones by automatically block-diagonalizing confusion matrices.

ACKNOWLEDGMENTS

iii

TABLE OF CONTENTS

Chapter 1

INTRODUCTION

1.1 Speech Perception

Speech perception happens naturally, long before an individual learns to read. The ability to discriminate and identify speech sounds is the common foundation for learning to read[1].

At least three channels are recognized in the reading and language comprehension process: the visual, auditory, and contextual channels (Allen, 2005; Kamhi and Catts, 2012; Bronkhorst *et al.*, 1993). The ability to recognize the image or shape of the letters and words (differentiating "q" from "p"or "b" from "d") contributes to the reading process and relies on the visual channel (Brandt and Rosen, 1980); the ability to sense and perceive the input sounds relies on the auditory channel; and the ability to understand concepts and meaning to aid in reading depends on the contextual perceptual channel.

For example, suppose, a subject was tested for speech perception with the word "speech", assuming the subject's vocabulary includes "speech", "speak" and "speed". If the subject didn't hear the last speech sound /tʃ/, the decision about which word was heard could still be narrowed down to the two words ("speech" and "speak"), with the help of information from the contextual channel. Additionally, if the subject didn't think this final speech sound sounded like a stop, the subject might choose "speech" as the answer. It is reasonable to infer, therefore, the decision was made with help from cues coming from both the contextual channel and

---

[1]From personal communication within research group.
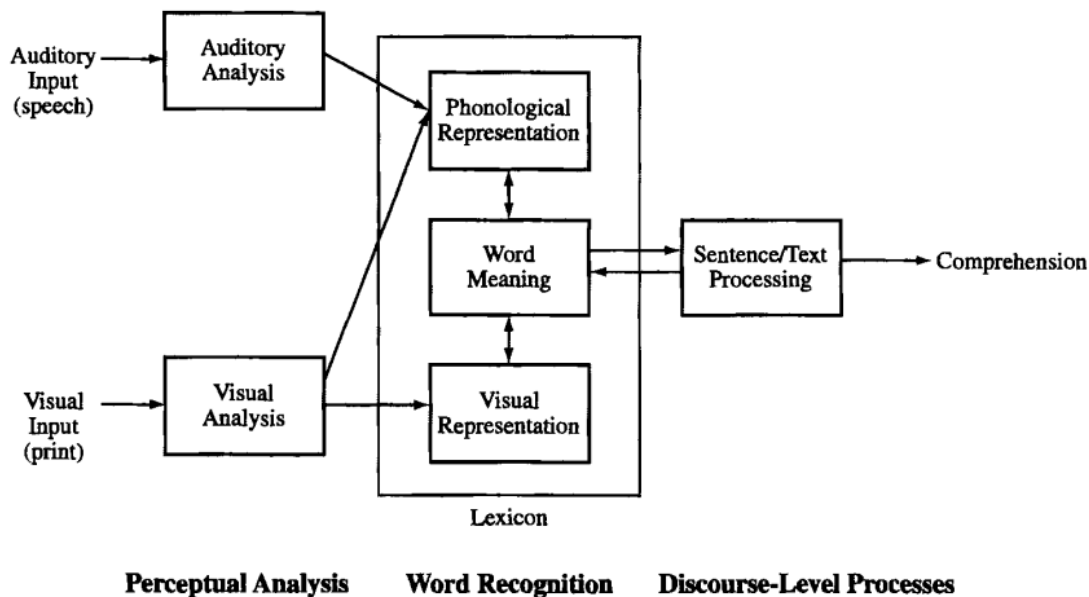
Figure 1.1: A model of spoken and written language comprehension. (Kamhi and Catts, 2012)

the auditory channel. Similarly, if a child has already acquired the grapheme or the written form of /tʃ/ (namely, "ch"), while an image of this phoneme was shown nearby at the same time during testing, unconsciously a hint might be formed, suggesting a /tʃ/ was heard. Only when all the interfering channels considered are effectively blocked (e.g., masked), can one truly measure the core speech perception ability of the subject.

There are two ways to measure speech perception: Articulation and Intelligibility (Allen, 2005). To properly introduce these two measuring systems, a couple of important concepts need to be defined first, to avoid potential confusions. The words **phone** and **phoneme** look deceptively similar. They differ. A **phone** is the smallest speech sound in the language of discussion (English in this). A phone could be a consonant or a vowel. However the concept of phoneme is based on its meaning. The smallest speech sound in a word (a sound with dictionary meaning) is a **phoneme**, like /b/, /ʊ/, and /k/ from the word "book". A **syllable** is a possible form of any number of consonants (C) and a vowel (V), such as V, CV, VC, CVC, CVCC, CCVC, and so on (eg. /paf/). To construct a maximum entropy (MaxEnt) syllable means to produce a syllable with a random phones from the a pool of all

2

possible phones in the language that is under discussion, so that each phone has an equal probability to go into the syllable. Such syllable is a MaxEnt syllable. The correct perception of recognition probability for a MaxEnt syllable is then the articulation score for this syllable.

The major difference between articulation and intelligibility is that articulation is the recognition of MaxEnt (nonsense) syllables, while intelligibility is the recognition of meaningful words (or part of meaningful words) (Fletcher, 1929). Articulation testing minimizes the contextual channel to aid in the phone recognition process. What it does require is the sensory capacity in the auditory channel to perceive phones without confusion, which in essence is phonetic perception (which will be introduced in 1.4). In other words, the difference between articulation and intelligibility finally comes down to the difference between phonetic perception and phonological awareness, or more precisely, the difference between phonetic perception and phonemic awareness (which will be introduced in 1.3). In addition, according to the summary by Share and Stanovich in 1995, compared with phonological sensitivity, contextual comprehension plays a less important role in reading comprehension or at least "decoding"in early reading (e.g. in first grade), which is the reason to focus more on phonetic perception while controlling and ruling out the contextual information when researching about possibly the most critical factors contributing to reading (Share, 1999; Share and Stanovich, 1995).

To summarize, the reading process could engage three channels: visual, auditory and contextual. However, the auditory channel in some instances is viewed as a more fundamental channel, that contributes the most to normal reading ability development, during the early developmental years. To study the contribution to reading ability coming from this channel, we must eliminate the contributions from visual aid and contextual aid. Therefore, graphs of letters or objects, and words (meaning) should be avoided when testing the auditory ability of subjects with reading disabilities (RD).

Table 1.1: Definitions of related concepts

Phone	smallest unit of speech sound
Phoneme	smallest unit of speech sound with meaning (word)
Syllable	a combination of any number of consonants (C) and a vowel (V), such as V, CV, VC and so on
Maximum Entropy (MaxEnt) Syllable	Syllable constructed by random choice of consonants and a vowel
Phonological Awareness	the multilinguistic skill of explicitly attending to, judging and manipulating sound structure in meaningful context, on syllable, onset-rime and phoneme level
Phonemic Awareness	the multilinguistic skill of explicitly attending to, judging and manipulating sound structure in meaningful context, only on phoneme level
Phonetic Perception	the auditory ability of hearing, perceiving and categorizing phone units, independent of meaning
Articulation	recognition of MaxEnt syllables
Intelligibility	recognition of meaningful words

1.2 Reading Disabilities (Kamhi and Catts, 2012)

Dyslexia, as a type of reading problem, specifically refers to the difficulties a subject has with word recognition (Kamhi and Catts, 2012). The International Dyslexia Association (IDA) characterizes this type of reading disability as inaccurate word recognition, and poor spelling and decoding. It is also popularly known as (specific) reading disability in most of the research literature. In modern medical diagnosis, subjects with dyslexia have neither outright brain damage nor sensory (hearing or visual) deficits, but still score low in tests with word reading. Word recognition holds a crucial place in the chain of spoken and written language comprehension development (the big central box of Figure 1.1). It correlates with performance for other reading abilities such as vocabulary, grammar, comprehension, reasoning, and so on. Therefore dyslexia is often equated to a more general term: **Reading Disability**. However, when children are tested to diagnose whether they have reading disabilities, they are usually tested for all abilities related to reading, not just word recognition. A phonological processing deficit has long been identified as the core of dyslexia. People consider the phonological deficit

to have a causal connection to dyslexia. Among all the phonological processing abilities, phonological awareness is discussed in comparison with the new concept to be defined in this report as phonetic perception.

1.3   Phonological and Phonemic Awareness

The connection to speech perception, viewed here as articulation, is typically regarded as intelligibility, while intelligibility plus reading ability are traditionally thought of as phonological awareness. There have been at least three decades of phonological awareness studies (Kamhi and Catts, 2012) (Goswami and Bryant, 1990) and the topic is still being researched (DeGroot *et al.*, 2015).

Phonological awareness is regarded as the multilinguistic skill of explicitly attending to, judging, and manipulating sound structure on the level of the syllable, onset-rime, and phoneme in language, that is, speech sounds used in a meaningful context. It is closely assumed to be related to cognitive function, which involves complicated top-down processes, that are difficult to control experimentally and thus to study.

The multilinguistic skill of attending to and manipulating sound structure on the level of phoneme in language is **phonemic awareness**. Both phonological and phonemic awareness are only related to spoken language from phonics, and refers to speech sounds that correspond to letters in print. Phonemic awareness is considered necessary for phonemic decoding ability (Kamhi and Catts, 2012). Numerous longitudinal studies and experiments have shown that, compared with rhyme awareness and verbal short-term memory, phonemic awareness has more predictive power for future reading abilities (Melby-Lervåg, 2012). A typical phonemic awareness assessment could include initial and final phoneme identification (/m/ in mat, /t/ in mat), phoneme blending (/f/+/æ/+/n/ = fan), phoneme deletion (Carlos ⇒ arlos), phoneme segmentation (cat = /k/+/æ/+/t/), phoneme synthesis (/m/+ onathan = Monathan), sound comparison (/d/ in good ≠ /k/ in book) and oddity

test (snake, lake, rake and pen) (Carson *et al.*, 2013). As pointed out in this review of the literature, phonemic awareness has great impact on the early development of reading ability, because it facilitates word decoding, with its corresponding speech sound recognition ability. Therefore, phonological processing deficits are assumed to be related to dyslexic reading problems.

Bradley and Bryant in 1983 published their monumental research on CVC real-word training for 65 children with reading difficulties (at least one standard deviation below the mean in reading level). The investigators pointed out that the RD group that received phonemic awareness training, with visual presentation, improved greatly in reading, compared with the group that did not receive such training. Because the 65 children were taken from a group of subjects from a previous test of reading ability on the same words, the experiment was longitudinally controlled. Hence the experiment presented one of the first demonstrations of a causal relation between phonological awareness and improvement in reading ability. Compared with the number of words used in most other sound categorization tasks, the list of CVC words used in this experiment was much larger, at least 72 different words ($3 \times 6 \times 4$). However, there were some shortcomings in the study. Firstly, not all of the vowels and consonants were used. Secondly, since meaningful words were used, the context channel of speech could contribute to or compensate greatly for the perception of the phones, even for children who lagged in reading. Thirdly, the sounds were produced in real time, so that other speech perceptual features such as tones and talker variation were not carefully controlled (Bradley and Bryant, 1978) (Bradley and Bryant, 1983).

In another published study (Manis *et al.*, 1997), 25 dyslexic (DYS) students (4th-10th grade), 25 chronologically age matched (CA) students (5th-8th grade) and 24 reading level matched (RL) students (2nd-3rd grade) were recruited to test the possible presentation of speech perception deficits in children with dyslexia. The spoken words /bath/ and /path/ were used as testing materials, to test the /b/-/p/ categorization ability of the groups of subjects. In general, the slope of the catego-

rization curves for the DYS group overall was shallower than those for the other groups, but the differences were not statistically significant. However, a subgroup of DYS students reported to have significantly poorer phonemic awareness.

Other experiments also failed to validate the existence of phonological or phonemic awareness deficits in children with reading disabilities. For example, a study (Blomert *et al.*, 2004) was carried out to investigate the role of the contextual channel in dyslexic children using spoken words. Participants were 42 graders (aged 7 to 10, 14 with dyslexia). Natural speech sounds were used in the words /tart/ and /kart/ to test for contextual sensitivity, and auditory, phonetic and phonological processing ability in children with RD. Two consonants, /t/ and /k/, in words were the targeted phonemes in these words. There were no substantial differences showing phoneme categorization or phonological deficits. With the help of context, children with reading difficulties performed at the same level as children with normal reading ability. These results were inconsistent with the initial assumption of the authors of a weak context sensitivity, and categorical-perception and auditory deficit in children with dyslexia.

Another study (Joanisse *et al.*, 2000) tested the relation between phonemic awareness deficits and reading disabilities. Researchers compared the performance of 61 dyslexic 3rd graders (aged 7-10), 52 chronologically age matched 3rd graders (aged 7-10) and 37 reading level matched 1st and 2nd graders (aged 6-8). Their categorical perception abilities were tested with meaningful word pairs: 'dug'- 'tug'and 'spy'- 'sky'. No significant difference in performance was found between the dyslexic group and the control groups.

In summary, it is difficult to analyze a speech perception problem using meaningful speech as testing materials. It is helpful to break the problem into parts, along its natural layers, and tackle them one after another. In these previously mentioned studies, phonological awareness and the cognitive level of processing ability were tested, to investigate the root cause of reading difficulties, but according to Harvey Fletcher's model, Figure 1.2 (Allen, 2005, Fig.3), these abilities lie
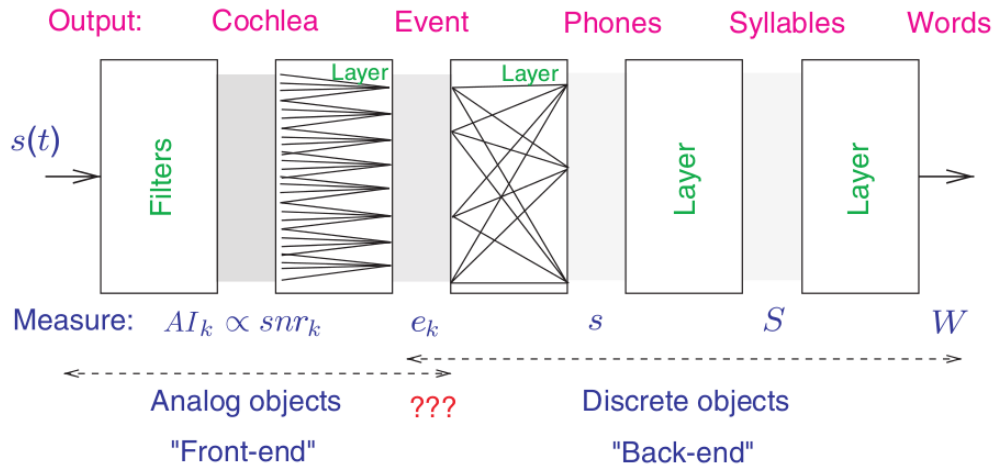
Figure 1.2: Model block diagram summary of human speech recognition. At the top of each block is a label that attempts to identify the physical process. The labels below the boxes indicate the probability measure defined at that level. See the text for the discussion of objects at the very bottom. The speech $s(t)$ enters on the left and is processed by the cochlea (first block), breaking the signal into a filtered continuum of band-passed responses. The output of the cochlea is characterized by the specific AI_k, a normalized SNR, expressed in dB units. The second box represents the work of the early auditory brain, which is responsible for the identification events in the speech signal, such as onset transients and the detection of basic measures. The third block puts these basic features together defining phones. The remaining blocks account for context processing (Allen, 2005).

in the very last stage in the speech recognition or speech perception process. In Fletcher's model, phones (the discrete speech units of the sounds of the language) become phonemes, phonemes are woven into meaningful contexts such as words, and words comprise reading materials. Therefore, relatively speaking, *phonetic* recognition would be a more basic (earlier) layer of speech processing. The most basic step in perceiving speech materials is based on phones. This perception of phones may well determine the success of reading and comprehension. Thus, the preceding analysis justifies our investigation of phonetic perception as a first step in the analysis of reading problems.

1.4   Phonetic Speech-Perception

By focusing on the perception of consonant-vowel (CV) and vowel-consonant (VC) *nonsense* syllables (speech units with maximum entropy), the present study explores the perception of the basic unit of the speech sound. In this way, the ability to perceive phones was decoupled from the lexical, syntactic, morphemic, and semantic information processing of speech context, making it possible to gain a more fundamental understanding of *phonetic* perception (speech perception) for children with reading difficulties.

Phonetic Perception, as described here, is the auditory ability of hearing, perceiving and categorizing phone units, independent of meaning. Such bottom-up processing ability is immediately available following birth. It is closely related to the sensory functionality of the cochlea, which transforms the acoustical vibration code into electrical signals, passed on to auditory neurons.

Following a review of various basic auditory skill studies on rapid auditory processing, synthetic syllable discrimination and categorization, gap detection (temporal resolving), formant transition and so on, Rosen in 2003 pointed out some association, but no causal relation, of auditory deficits (temporal processing or other) to specific reading disabilities (Rosen, 2003). He also pointed out the lack of any conclusions by these studies that there is a relation between nonspeech, auditory perception deficits and speech perception deficits. In another study (Brandt and Rosen, 1980), 12 RD and 4 boys who served as reading control subjects (RC) were tested with synthetic CV syllables for auditory perception deficits. The consonants used in the test were all stop consonants and there were only five in total: /d/,/t/,/b/,/g/,/k/. All were presented in the context of the vowel /ɑ/. This study concluded that there was no significant impairment in phone perception in children with reading disabilities[2].

In another study, intervocalic consonant discrimination was examined in the

---

[2]Recently a study also used /dɑ/ as testing material to predict language or learning disabilities for children (ASHA, 2015).

form of VCV structure (Hazan and Adlard, 1998). The 15 consonants tested were /b/, /d/, /g/, /k/, /v/, /ʒ/, /ʃ/, /s/, /f/, /m/, /n/,/r/, /w/, /j/, /l/ with the vowel /ɑ/. Again there was no general group difference found in this consonant perception task for reading disabled (RD) and reading control (RC) children. Post-hoc investigation with ANOVA revealed that only a subgroup of RD children had significant deficits compared with others. For this subgroup of RD children, phonological awareness was also significantly worse than that for others in the RD group. Still, in later literature (Hazan *et al.*, 2013), these authors concluded for this study that no significant consonant categorical deficit was found for RD children.

Messaoud-Galusi and colleagues in 2011 carried out intricate consonant identification and discrimination experiments on 62 DYS and 51 CA children (aged 6-14), to investigate the conjecture of impaired speech perception abilities in children with reading disabilities (Messaoud-Galusi *et al.*, 2011) . The experiments were carried out in quiet and in 20-talker babble noise, with only two word-like CV syllables BEE and PEA. The authors concluded that there was no consistent speech perception deficit in dyslexic children. Two interesting results were noticed: First, the investigators reported better performance for RDs in noise than in quiet, and second, age did not have an effect on within-category discrimination for children with dyslexia. This experiment was a follow-up investigation of a similar study done with adults, which used the same syllables and procedures and also suggested weak support for a speech perception deficit in dyslexia (Hazan *et al.*, 2009).

Another study on 23 DYS and 22 RC children (aged 8-12) (White *et al.*, 2006) concluded that there is little correlation between dyslexia and an auditory sensorimotor impairment. This study used a /ba/ to /ga/ continuum and asked subjects to indicate which syllable they heard (in a categorization perception task). In contrast to the preceding study in the present reviews, the authors reported a significant impact of speech perception on other phonological skills. Unlike studies which reported no significant correlation between speech perception and reading

disabilities, in a study with 19 DYS (aged 8-12), 18 CA (aged 8-12) and 19 RA children (aged 6-8), speech perception deficits in dyslexics were reported while tested in noise, but not in quiet, with VCV nonsense syllables with the vowel /ɑ/. The consonant tested was one of 16 possible French consonants (/p/, /t/, /k/, /b/, /d/, /g/, /f/, /s/, /ʃ/, /m/, /n/, /ʀ/, /l/, /v/, /z/, /j/) (Ziegler *et al.*, 2009).

To sum up, the relation between speech perception and reading ability development is highly variable. Therefore a thorough investigation on this relation is in order, to clarify the roles of the various processing channels and evaluate the classic Fletcher model in Figure 1.2.

1.5 Development of Reading Abilities and the Importance of Correct Phone Perception in Early Childhood

The development of reading ability is closely related to language, learning and every literate aspect of our lives. While the ability to comprehend is closely related to background knowledge, the decoding ability is relatively independent in terms of knowledge domain. There are those who tend to believe that speech perceptional errors are due to genetic mis-wiring of cortical areas, for example. If this is the case, plasticity can play no role in the relearning. Our alternative hypothesis is that plasticity can and does play a role and that the problem is due to improper learning of some basic speech sounds in early development. It is widely acknowledged that middle ear hearing loss can cause hearing threshold elevations of above 50 dB. For example, if you have hearing loss for a couple of months, and if this happens during in a critical period of learning, you can miss the correct decoding of phones. How children learn speech perception is one of the mysteries in communication development. One hypothesis is that accidental or consistent exposure to false representations of speech phonetics in early childhood results in speech perception errors and consequential reading difficulties. Training could help alleviate the temporal processing deficit of language impairments (Merzenich and

Jenkins, 1996).

The accepted belief, that the left hemisphere is the controlling hemisphere for speech and language, may not hold true in children with Autism Spectrum Disorder (ASD). In a similar way, ASD could play a role in a delay in consonant and vowel acquisition (Chenausky, 2015). Except for children with communication impairments like those found in ASD, hearing impairment, or cognitive deficits, who might need to receive intervention later, generally the timing for speech and language perception intervention should be as early as kindergarten or even earlier. Speech-language pathologists do "early intervention" for children from birth to three years old. They also help children with phonology all the way into high school. The chart of normal speech and language perception and production development for early childhood is reproduced here as Figure 1.3 (Kuhl, 2004, Fig.1) (Chall, 1983).

Effective reading comes from successful word recognition and is usually thought to be fundamentally rooted in phonemic decoding. Describing basic sound-letter correspondence, phonemic decoding ability requires delicate and precise auditory and acoustic knowledge of speech sounds. However, word recognition could be significantly more complex than phonemic awareness (Figure 1.2). After achieving phonemic awareness, children would start to form spelling patterns of letters in chunks, to store in long-term memory. Then they would begin storing spelling patterns, decoding the words, memorizing them after repetitions, followed by building up a sight vocabulary of whole word. Beyond this point, the recognition process becomes automatic and no longer requires too much decoding effort. This completes the early stages of learning to read. The Self-Teaching Hypothesis claims the reverse is true (Share and Stanovich, 1995). After the maturing of phonemic decoding, the ability of making analogies to known words would further enhance reading ability. Children with normally developing language skills start to produce their first word in their mother tongue at about 12 months (Figure 1.3),
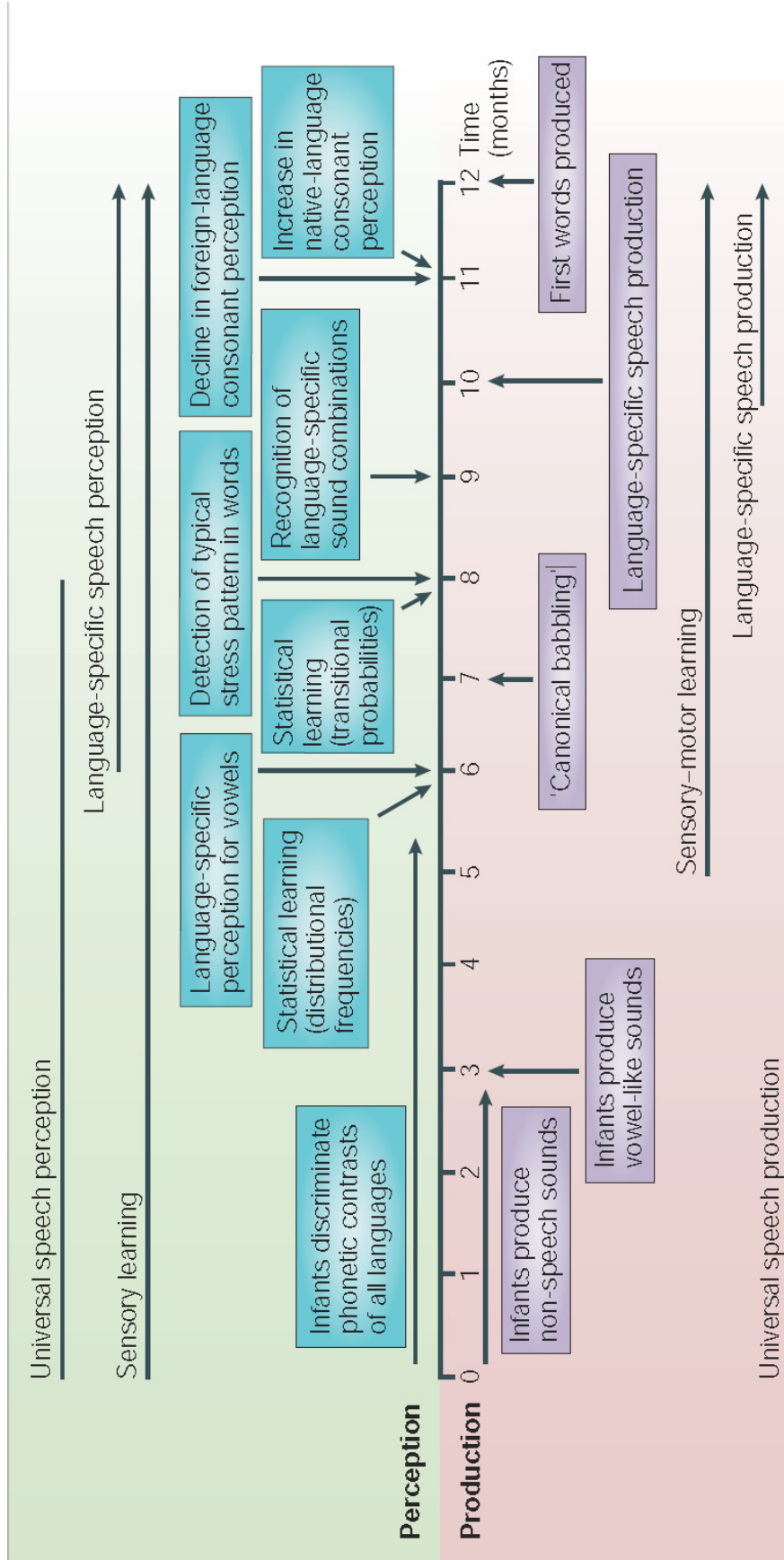
Figure 1.3: Speech perception and Production development during early childhood (Kuhl, 2004).

and they should have mastered phonemic awareness and word recognition by the third grade.

Research shows that critical words contributing to the meaning of a passage can only be contextually guessed at a correct rate of 29.5% for some materials (Share, 1995). Relying only on context to guess words is not an accurate way to read, but it allows poor readers to use context as a major source for word identification, while skilled readers do not have that dependency (Oakhill and Beard, 1999). As said, reading ability is gained in steps. Although this kind of complex ability could be developed and improved throughout one's entire life, the optimal period is early in childhood (Kamhi and Catts, 2012). It cannot be stressed enough that it is easier to prevent than to alleviate reading disabilities. Early testing and training on phoneme awareness should be strongly promoted, to prevent later appearing reading difficulties (Melby-Lervåg, 2012). However, phonemic awareness development involves information contributed by the contextual channel. If phonemic awareness training fails to improve reading abilities, the possibility exists that the problem is at the phone level (Figure 1.2). By focusing only on the auditory channel, we might be able to understand the role of speech perception in reading and eventually establish a training program that effectively contributes to reading development.

1.6   Research Questions

In general, the relation between the three concepts introduced – Phonological Awareness, Phonemic Awareness and Speech Perception (includes phonetic perception) – is depicted in Figure 1.4. Being the foundation of normal speech communication, speech perception shares much in common with the hearing process (Paul and Whitelaw, 2011). Before the higher level cognitive processes (such as comprehension and reasoning) can develop, the lowest level of cognitive function, phonemic awareness, must function. The brain first recognizes the sounds
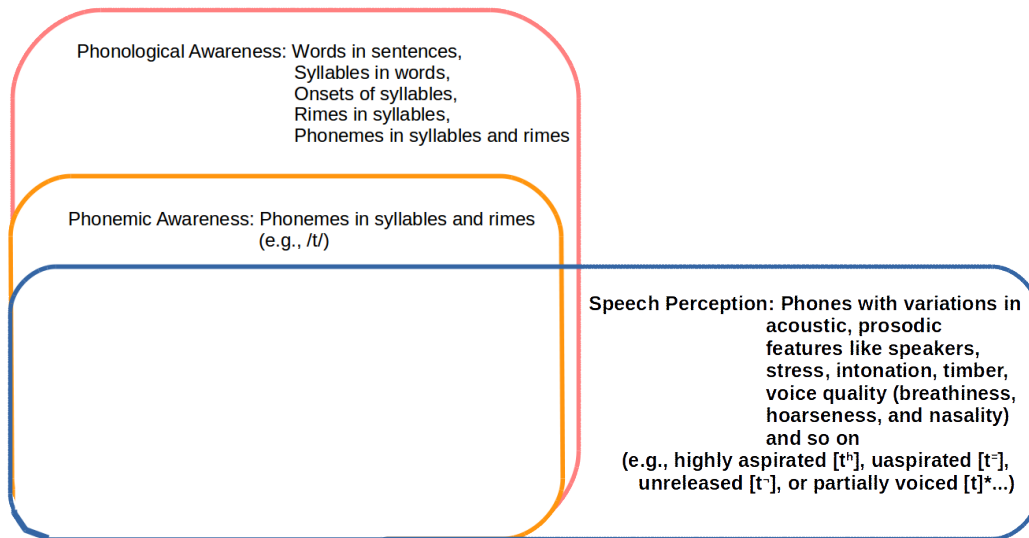
Phonological Awareness: Words in sentences,
Syllables in words,
Onsets of syllables,
Rimes in syllables,
Phonemes in syllables and rimes

Phonemic Awareness: Phonemes in syllables and rimes
(e.g., /t/)

Speech Perception: Phones with variations in
acoustic, prosodic
features like speakers,
stress, intonation, timber,
voice quality (breathiness,
hoarseness, and nasality)
and so on
(e.g., highly aspirated [tʰ], uaspirated [t˭],
unreleased [t̚], or partially voiced [t]*...)

Figure 1.4: Relation between the three kinds of capabilities: Phonological Awareness, Phonetic Awareness and Speech Perception.

on the phone level, then it combines more information to form syllables, based on the sounds located before or after the sound being processed. Starting from here, other cognitive top-down abilities would kick in to reconstruct a clear and solid abstract concept of the combined sounds. Later, the recognition is transformed into graphemes and the hearing process ends. At this stage, auditory memory is crucial. Visual recognition of graphemes in words is the following stage in reading.

As discussed in the previous section, even though a lack of phonemic awareness is viewed as a causal factor in reading disabilities, there are still studies showing the two are uncorrelated. A more basic capability, that supports early reading ability development, is phonetic perception. Few findings have been reported to map out the phonetic perception of children with reading difficulties. The perceptual limits for various phones for reading disabilities remains unknown; however, Fletcher's 1929 model of speech perception seems relevant (Figure 1.2). Hence, in this study, the following research questions were explored:

1. Without access to visual and contextual information, do children with read-

ing disabilities encounter a phonetic perception deficit?

2. If they do, what kind of phonetic perception task is more effective in diagnosing their reading disability? What kind of phonetic perception task is more informative about their weakness regarding specific phones?

3. Do children with reading disabilities and normal children have both common and unique phonetic perception patterns?

1.7   The Design of the Perception Tasks in This Study

The first task of this study is the syllable confusion oddball task. Children in this task listen to three natural speech syllables in each trial and are expected to pick out the odd sound which contained a different sound structure than the other two. Similar oddball tasks were performed before (Bradley and Bryant, 1978) (Bradley and Bryant, 1983) but with real words. In these studies, a causal relation between phonological awareness and improvement in reading abilities was found. However, as discussed before, the usage of real words as testing materials is considered contextually indicative. Children who have the opportunity to be exposed in certain literacy environments could have an advantage in listening comprehension for words that presented in this kind of oddball task. In the oddball task proposed in this study, the testing materials are all purposefully nonsense syllables. These syllables have no meaning in English, but they all sound like English. We propose that using materials with no literal meaning could isolate the acoustical perceptual channel from the contextual channel and, hence accomplish the goal of testing only the sensory perceptual ability of the children.

The second task – the nonsense syllable confusion matrix – is specially designed to obtain the intangible perception of phones in each child. It also uses nonsense syllables as testing materials so as to block the contextual channel. Another advantage of this task is that it is assumed to require no other cognitive processing ability and therefore requires less effort.

1.8   Special Advantages of Both Phonetic Perception Tasks

There were four main differences between the design of this experiment and that of speech perception tests. The sounds were pre-recorded and selected to remove any real time or loudness variation between phones. Both tasks had extensive coverage of consonants and vowels in English: 24 consonants in total and 15 vowels in total. The sounds were produced by 18 talkers to better represent the variation encountered with speakers of English. Instead of words, maximum entropy syllables were used as testing materials. This could guarantee the isolation of information to be supplied by the auditory channel, reducing the interference from the contextual channel. Because the children only needed to respond with the original information in the presented tokens, there was no need for the children to segment phones from the syllables or to manipulate them. Hence the task was much easier than a phonemic awareness task. There was also no influence of a visual channel. In our experiment, the children had to rely exclusively on the auditory channel, for input.

Chapter 2

METHODS

There are two phonetic perception tasks involved in this study: the Syllable Confusion Oddball (SCO) Task, a speech perception discrimination task; and the Nonsense Syllable Confusion Matrix (NSCM) Task, a speech identification and production task. During each task, the children were given game breaks (five minutes of break for every ten minutes of testing), and enough rest and treats, to avoid possible fatigue and boredom. Each child participated in the experiments for up to 10 weeks of two hours per week. Each week they took one hour testing of the SCO task and one hour testing of the NSCM task. Totally on average each child completed 1600 trials of the SCO task, with a standard deviation of 548 trials.

2.1 Participants

All the children with RD recruited in this study came from the same weekly reading group program in Champaign, Illinois. They all passed a pure-tone hearing screenings (500Hz, 1kHz, 2kHz, and 4kHz in each ear, at 20 dB SPL), indicating normal hearing ability. There were no visual, neurological, cognitive or emotional problems reported. All the children had normal or above normal IQ, and were 8 to 11 years old. This is a typical age range where deficits are discovered, but not overcome, during reading development. The reading disabled (RD) group had nine children (six girls), all aged eight to ten years. The reading control (RC) group had six children (two girls), aged eight to eleven years. To assess their reading abilities, a battery of reading tests were administered. These tests were the Woodcock

Reading Mastery Tests-Revised (WRMT-R), specifically the Word Identification subtest (WI) and Word Attack subtest (WA); and Gray Oral Reading Test in 4th edition (GORT-4), which included a Fluency score (R-FLU) and Comprehension score (R-COMP) (Johnson *et al.*, 2015). For each measure, the children in the RD group scored significantly lower than the children in the RC group. The t-values between group scores were 4.34, 5.18, 4.03 and 4.64 for WI, WA, R-FLU and R-COMP respectively. The p-values were 0.0017, 0.00029, 0.0048 and 0.00049. For all four reading measures, the RD group scored at least one standard deviation below normal reading (RC) children according to the criterion of each measure. Detailed scores are listed in Appendix B.

2.2   Preparation

Natural speech sounds have more perceptual cues than synthetic sounds, thus are considered better for human speech perception tests (Li, 2009). The set of natural sounds that were used for both tasks came from the commercial Linguistic Data Consortium (LDC-2005S22) database (Fousek *et al.*, 2004). It consists of all diphone syllables allowed in English (i.e., CV and VC syllables) formed with 24 consonants and 15 vowels, spoken by 18 different talkers. The consonants are /p/, /b/, /d/, /t/, /g/, /k/, /m/, /n/, /l/, /f/, /v/, /s/, /z/, /ʃ/, /ʒ/, /r/, /ʤ/, /tʃ/, /θ/, /ð/, /ŋ/, /h/, /w/, and /j/. The vowels are /ɔɪ/, /aʊ/, /eɪ/, /aɪ/, /i/, /ɝ/, /u/, /oʊ/, /ɪ/, /ɔ/, /ʊ/, /ʌ/, /æ/, /ɛ/, and /ɑ/.[1] Some syllables are accidentally real monosyllabic words but in most cases they are only nonsense (MaxEnt) syllables. All sounds were played without noise, in a random sequence, at the most comfortable loudness level for each child. The sounds could be replayed as many times as needed, but few sounds were requested for more than three presentations. All children listened to the sounds through AKG K240 Monitor headphones (circumaural, made in Austria,

---

[1]Instead of International Phonetic Alphabet (IPA), some figures in this report were generated with Darpabet to represent the phones. Therefore the conversion between Darpabet and IPA was presented in Appendix A.

600 ohms) from laptops at a comfortable loudness level chosen by themselves at the beginning of the session.

2.3   The Syllable Confusion Oddball (SCO) Task

In this task, a combination of three random nonsense (maximum entropy) syllable sounds (either all CVs or all VCs) were played to the children. Two of the syllables contained the same sounds spoken by two different talkers. The remaining syllable was a different sound spoken by a third talker, that differed in only either C or V. The children were asked to point at one of three wooden blocks (labeled with the numbers 1,2, or 3), corresponding to the oddball sound. The response was recorded by accompanying staff members. For example, suppose the string of syllables to be played is ta - da - ta. It is a set of syllables with consonants in the initial position, and with the vowels in the final position. In this example, the second syllable is the oddball. The oddball sound could be randomly placed in any position (i.e., CV syllables). In this example, the task was to identify the 2nd block on the desk as a correct response. If the child pointed to the 1st or 3rd block, that would be an error. Response data were accumulated to calculate the probabilities of error, used in the analysis. The average number of trials for consonants in initial position for all children was 42, with a standard deviation of 19; vowels in initial position 43, with a standard deviation of 19; consonants in final position 44, with a standard deviation of 17; and vowels in final position 45, with a standard deviation of 15.

2.4   The Nonsense Syllable Confusion Matrix (NSCM) Task

In this task, the children were tested on their ability to perform in a way that required both speech identification and speech production. In each trial, a single

random MaxEnt syllable was played to the child. He or she was asked to repeat the syllable afterwards. For example, when syllable 'at' is played, the response from the child would be evaluated by two transcribers (judges). If the judges agreed on hearing a 'ak' as the response, then 'ak' would be entered into the system. If the judges disagreed, both responses were entered, with a weight of $\frac{1}{2}$. The judges would not be informed about the true speech sounds before they entered the child's response. For this task, the average number of trials for consonants in syllable-initial position was 31, with a standard deviation of 3; the average number of trials for vowels in syllable-initial position was 42, with a standard deviation of 3; the average number of trials for consonants in syllable-final position is 30, with a standard deviation of 5; and the average number of trials for vowels in syllable-final position was 47, with a standard deviation of 4.

2.5 Statistics Tests

The data of our perception tasks is processed with arcsine transform before the quick ANOVA check. The equation used for rationalized arcsine transform is $pAsin = 46.4732 \times 2 \times asin(\sqrt{p}) - 23$. Studebaker in 1985 deducted the constants in this equation for rationalized acsine transform used in speech research data processing (Studebaker, 1985) (McDonald, 2014).

In this thesis, we first considered using ANOVA as a testing method to quickly identify significant factors or interactions in our experiment. After the quick testing of these factors, we also fitted generalized linear mixed effect regression models to the data. Although it requires more computation, regression is more generalized and has fewer restrictions than ANOVA. Detailed explanation and discussion will be given in the following chapter.

To reveal the differences in performance for the two groups of children, the sorted error patterns were plotted for each child in solid lines. Here is how we plotted the sorted error graph: The x-axis represents the phones sorted by error.

So the first tick on x-axis represents the phone with least error, and the second tick represents the phone with second least error. The y-axis displays the values of probability of error for the phone. The x-axis could vary a great deal from subject to subject, as it did for Alina and Bob in Figure 3.1.

The RDs' performance in the following sorted error plots were emphasized using colors, contrasting with the gray curves to represent the control group. The y-axis data are presented in log percent error. Only the pseudonyms for RDs are displayed in the legend.

Chapter 3

RESULTS AND DATA ANALYSIS

3.1 Sorted Error Plots for SCO Task

3.1.1 Consonants in syllable-initial (CI) position

In the top left panel of Figure 3.2, gray curves were for the sorted probabilities of error responses to all the consonants in syllable initial position from the RC group[1]. Probabilities of error ranged from 2% to 30%. Meanwhile probabilities of error for the RD group generally ranged roughly from 7% to 50%, with two obvious exceptions: RD-Teddy and RD-Norene, who did better than the rest of the RD group. Results from the other 3 panels of Figure 3.2 revealed similar characteristics for both the RD and RC children.

The top left panel of Figure 3.2 revealed a striking separation of the RC and RD children when discriminating consonants in the syllable-initial position. A huge gap between the colored and the gray curves distinguished the upper group from the lower group, indicating that children in the upper group (RD-Alina, RD-Angela, RD-Edward, RD-Latisha, RD-Laura, RD-Shauna, and RD-Tony) are struggling with discriminating all consonants, starting with a lowest 7% error. The six of them were all profiled as children with reading disabilities.

RC-Anton, RC-Evan, RC-Carly, RD-Teddy, RD-Norene, RC-Miguel, and RC-Bob, however, had a better performance in discriminating consonants in syllable initial position than most of the children in the RD group, that is, they all had

---

[1]Data for RC-Joanna were not presented in the plot, because the total number of trials she did was less than ten for some phones.
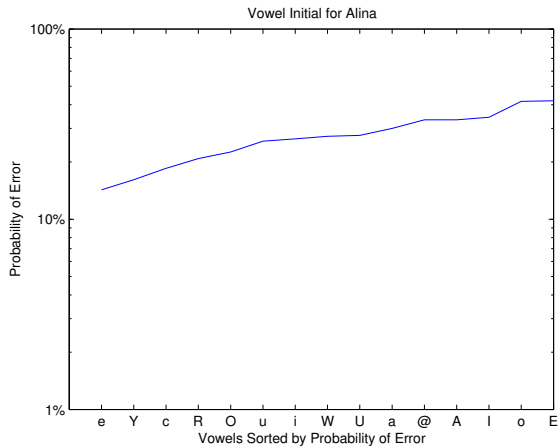
(a) Sorted error in VI for Alina (b) Sorted error in VI for Bob
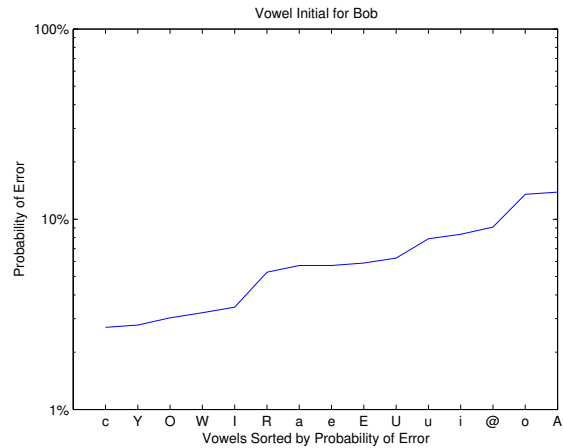
Figure 3.1: Sorted error plot examples: The phones in the x-axis were sorted by their probabilities of error. Therefore the sequence of phones arranged in the x-axis can vary from subject to subject.

lower sorted probability of error. RC-Anton had no error for 10 out of the total 24 consonants. RC-Carly also had zero error for 7 out of the total 24 consonants. The sorted error curves for RC-Anton, RC-Evan and RC-Carly were the lowest among all, with probabilities of error values ranging from less than 2% to less than 20%. All three of them are RC.

RC-Bob and RC-Miguel with normal reading ability were at the same level as RD-Norene and RD-Teddy. The probabilities of error for them ranged from 3% to 30%. We could easily spot that RD-Laura and RD-Edward had the largest slopes, meaning that some consonants were much more difficult for them to perceive than other consonants. Between them, RD-Laura had the smaller intercept in the y axis, meaning smaller probabilities for the lower ranking phones. All RCs and RD-Norene and RD-Teddy in RD had smaller slopes, with smaller intercepts on the y axis, indicating relatively uniform performance for all consonants and overall smaller error rates.
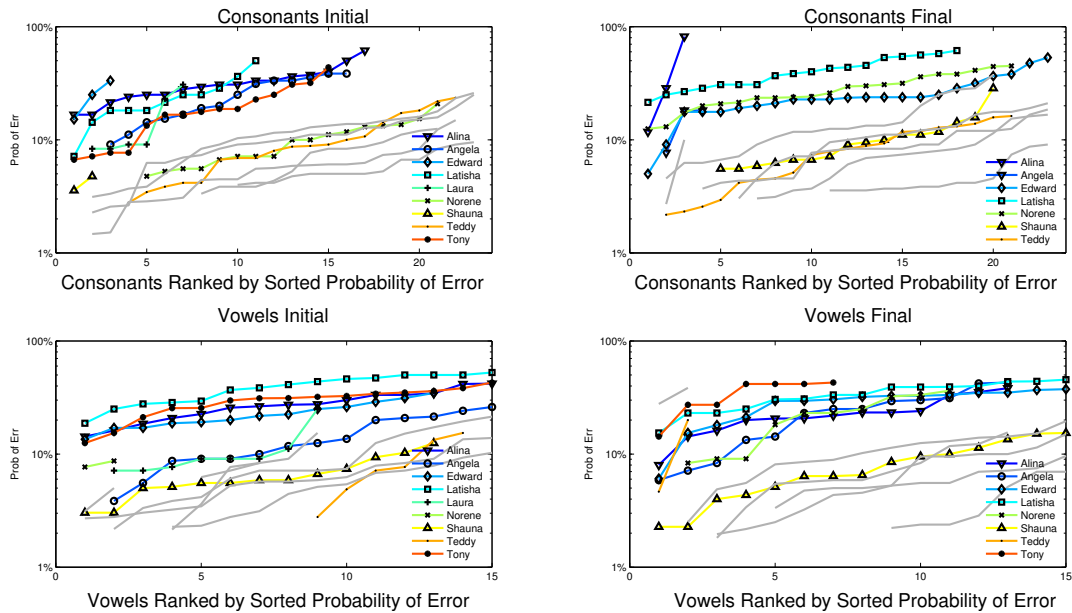
Figure 3.2: Sorted Error Plots for SCO task. That the curves do not start at 1 on x-axis indicates zero error for those early-ranking phones; and that the curves do not end at 24 (consonants) or 15 (vowels) indicates 100% error for those late-ranking phones. The curves tend to have greater slopes for the first five phones in each case. The slopes get smaller after that but the probabilities of error stay high.

3.1.2 Consonants in syllable-final (CF) position

The top right panel of Figure 3.2 presents the performance of all the children for consonant final position.[2] Although the resulting curves are not as well separated as for CI position, there is a clear tendency for RDs to have higher probabilities of error.

RD-Alina had more than 10% error for almost all consonants in final position. RD-Latisha, RD-Norene, RD-Edward and RD-Angela are also still on the higher error side.

RC-Anton still had no trouble discriminating 10 out of 24 consonants, with his highest probability of error for these sounds of less than 10%. RC-Joanna, RC-Evan and RD-Teddy were performing with less than 20% error. RC-Joanna had no error for 7 out of 24 consonants. The highest error rate for RC-Bob was slightly

---

[2]Data for RD-Tony and RD-Laura were not presented in the plot, because the total number of trials they did was less than ten for some phones.

greater than 20%. RC-Miguel and RD-Shauna had a little bit higher error, ranging from 5 to 30%.

Data collected from RC-Carly in CF strongly differed from her CI error. She had a minimum probability of error of around 3%, while for consonants at syllable initial position she had zero error for 7 out of 24. Data for RC children are presented in Appendix H.

We can also easily confirm that RC-Carly, RD-Alina and RD-Angela had the largest slope, meaning more abrupt increases in perceptual difficulties across consonants than other children. Between them, RC-Carly has the smallest y intercept. All RCs and RD-Shauna and RD-Teddy have smaller slopes, with smaller y intercepts, meaning that RCs and RD-Shauna and RD-Teddy had relatively uniform smaller probabilities of error for all consonants. RD-Latisha, RD-Norene and RD-Edward all had small slopes but with slightly larger y intercepts.

3.1.3 Vowels in syllable-initial (VI) position

The bottom left panel of Figure 3.2 is the resulting plot for vowels in syllable initial position situation. Previous high performers like RC-Anton and RD-Teddy still outperformed everyone, having between 2% and 20% error. RC-Bob in RC improved in his discrimination of vowels in initial position, compared with his discrimination of consonants in initial position.

RD-Alina had a different score range for vowels in initial position, which was narrowed to only 10 to 40%. RD-Latisha struggled the most with discriminating the vowels in syllable-initial position. Her probabilities of error for the target phones were 20% or greater. Besides RD-Alina and RD-Latisha, RD-Tony and RD-Edward were among the worst performers, with probabilities of error ranging from 10% to slightly higher than 40%; while RD-Angela, RD-Laura and RD-Norene, along with RC-Joanna, RC-Evan, RC-Carly, and RC-Miguel, were in the middle region, with probabilities of error ranging from 2% to 30%.

We could easily spot that RC-Carly, RC-Evan, RC-Joanna, as well as RD-Teddy and RD-Laura had the largest slopes, meaning more difficult perception for certain vowels than others. Among them, RC-Evan, RC-Joanna and RD-Teddy all had smaller y intercepts. Everyone else had smaller slopes, with larger y intercepts, meaning that they had uniformly large error for all vowels.

3.1.4 Vowels in syllable-final (VF) position

Similar to CI, the VF case also has a clear separation between the RD and RC groups.[3] Again most of the RD children (with the exception of RD-Shauna) fell as a cluster in the upper left corner of the panel in Figure 3.2, with high error rates for all vowels, while most of the RC children (with the exception of RC-Miguel) had relatively low probabilities of error. As in the other cases, RC-Anton was again the best performer of all in this case. He had zero error for 8 of the 15 vowels and his highest error rate was 7%. RC-Joanna was the second best performer, with her maximum error for all vowels being less than 10%. Data across all conditions showed that her phone discrimination ability was better when the phone is placed in syllable-final position, whether the phone is a consonant or a vowel. The rest of the lower error performers were RC-Evan, RC-Carly, RC-Bob, and RD-Shauna.

RC-Miguel in the RC group had the highest error rate among all children, starting at 30%. The rest of the high error group were RD-Tony, RD-Teddy, RD-Latisha, RD-Edward, RD-Alina, RD-Angela, and RD-Norene.

We can easily see that children RC-Anton, RC-Miguel and RD-Teddy had the largest slopes, meaning special difficulty in perceiving certain vowels. Among them, RC-Anton had the smallest y intercept. Everyone else had smaller slopes, with larger y intercepts, meaning that they had uniformly large errors for all vowels.

---

[3]Data for RD-Laura were not presented in the plot, because the total number of trials she did was less than ten for some phones.

3.1.5 Benchmarks generated from the RC group

On average, RCs have 6.44% error for consonants, 6.73% error for vowels, and 6.56% for all the phones. A profile of a hypothetical "normal" control subject is created from the average performance of the group, as presented in Figure 3.3. From Figure 3.3, we can see that in log scale with base 10, the variances for all four conditions (namely cf, ci, vf and vi) stabilized. In other words, the variances of the sorted errors in perceiving the phones increased exponentially from low error phones to high error phones. Individual RCs performances relative to the performance of this hypothetical average control subject are plotted in Figure 3.4. In Figure 3.5, histograms of score ratios, of individual RCs over the mean RC, are shown. The ratios of individual error rates to average error rates ranged from zero to two. The median for average performance in CF is 5.84%; that in CI is 6.99%; that in VF is 7.51%; and that in VI is 5.67%. The standard deviation from mean for CF is 5.76%; that for CI is 5.65%; that for VF is 7.06%; and that for VI is 4.98%.

3.1.6 Probability of error ratios and their distributions

On average, RDs have 20.31% error for consonants, 21.25% error for vowels, and 20.77% for all the phones. Probability of error ratios of RDs to the average RC level and the distribution of the ratios are presented in Figure 3.6 and Figure 3.7. For cases in CI, the probabilities of error from RDs were 4 to 5 times higher than the average RC probability. The only exceptions were RD-Norene and RD-Teddy, who had probabilities of error at the same level as the average RC probability. For cases in CF and VF, the exceptions were RD-Shauna and RD-Teddy. The ratios of probabilities of error from RDs to the average RC level for VI distributed exponentially from 0 to 10, with a mean of 2.7.
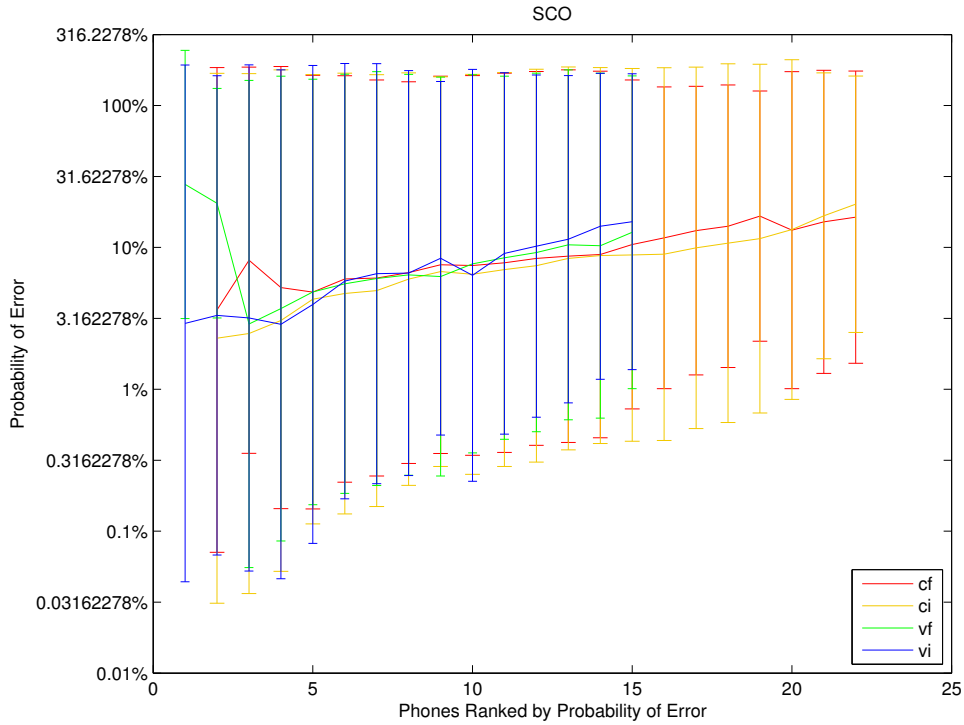
Figure 3.3: Average RC performance for SCO (log scale). The average RC erred at 6.56% for all phones with stabilized or slightly decreased variance

3.2 Repeated Measure ANOVA: Result for SCO

From the result listed in Table 3.1 on page 30[4], the ANOVA revealed a main effect of reading disability ($F(1,10)=17.62$, $p=0.00184$), with more errors for the reading disabled listeners than for the control listeners. There was also a main effect of phone position ($F(1,10)=6.67$, $p=.0273$), with more errors for final position than for initial position. None of the other effects or interactions were significant.

The mean values of all the main effects and significant interactions for Table 3.1 are listed in appendix D.

---

[4]All of the statistic tests in this thesis are performed with R. The probabilities of error were all arcsine transformed before the analysis of variance.

Table 3.1: Repeated Measure ANOVA Result for SCO task. Probabilities were processed by arcsine transform. Generalized Linear Model: ProbOfErrAsin $\sim$ SubGrp $\times$ PhonePos $\times$ PhoneType + Error(Individual/(PhonePos $\times$ PhoneType)).

Factor	Df	Sum Sq	F value	Pr(>F)
Between Individuals				
SubGrp	1	6655	17.62	0.00184
Residuals	10	3777	378	
Within Individuals				
(Between SubGrp and PhonePos)				
PhonePos	1	292.7	6.669	0.0273
SubGrp: PhonePos	1	14.5	0.331	0.5775
Residuals	10	438.9		
(Between SubjGrp and PhoneType)				
PhoneType	1	41.9	0.742	0.409
SubGrp: PhoneType	1	7.8	0.139	0.717
Residuals	10	565.5		
(Interactions)				
PhonePos: PhoneType	1	0.2	0.005	0.944
SubGrp:PhonePos:PhoneType	1	0.1	0.003	0.959
Residuals	10	323.6		

Table 3.2: Post Hoc Repeated Measure ANOVA Result for SCO task: RD-Norene, RD-Shauna and RD-Teddy were deleted from the analysis. Probabilities were processed by arcsine transform. Probabilities were processed by arcsine transform. Generalized Linear Model: ProbOfErrAsin $\sim$ SubGrp $\times$ PhonePos $\times$ PhoneType + Error(Individual/(PhonePos $\times$ PhoneType)).

Factor	Df	Sum Sq	F value	Pr(>F)
Between Individuals				
SubGrp	1	8629	52.99	8.56e-05
Residuals	8	1303		
Within Individuals				
(Between SubGrp and PhonePos)				
PhonePos	1	109.62	4.811	0.0596
SubGrp:PhonePos	1	1.06	0.046	0.8348
Residuals	8	182.30	22.79	
(Between SubjGrp and PhoneType)				
PhoneType	1	24.4	0.347	0.572
SubGrp:PhoneType	1	4.2	0.060	0.812
Residuals	8	563.4		
(Interactions)				
PhonePos:PhoneType	1	0.99	0.026	0.876
SubGrp:PhonePos:PhoneType	1	0.89	0.023	0.882
Residuals	8	305.25		
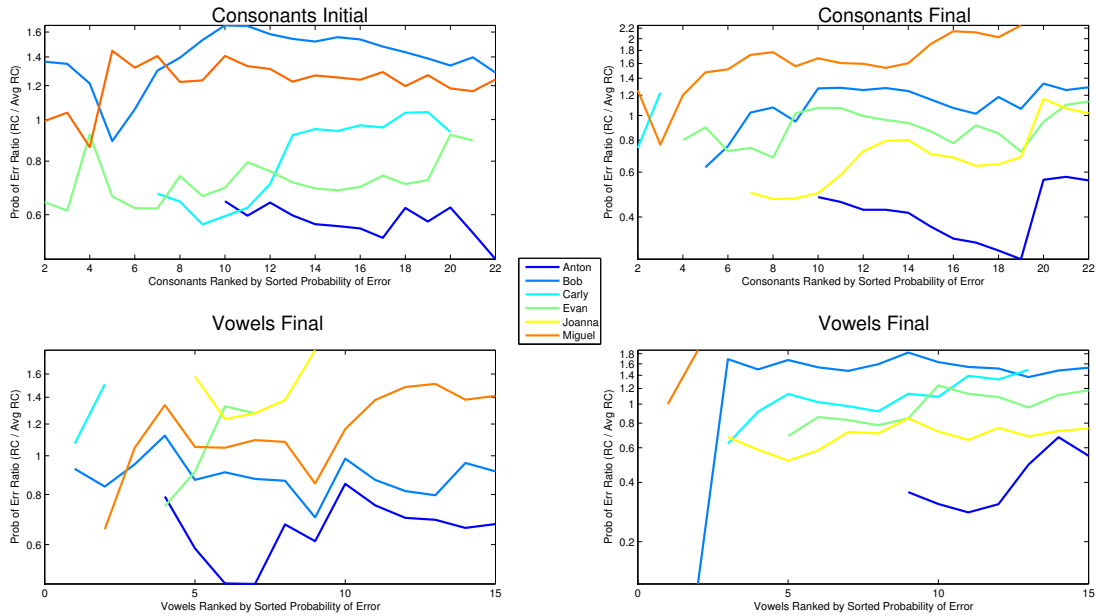
Figure 3.4: Probability of error ratios of individual RC to average RC for SCO. The probabilities of errors from each RC were divided by the probabilities of error from the average RC, and the resulting ratios were plotted in colored curves. Anton was the best RC among all children. Joanne perceived final phones (consonants and vowels) better than average RC. Miguel perceived weak in all four conditions.

3.3   Regression Analysis for the SCO Task

ANOVA requires the data set be balanced, that is, to consist of approximately normally distributed variables of nearly equal variance. However, regression analysis is not sensitive to these restrictions. Thus regression can function as a useful control on the ANOVA results. We fitted logistic mixed-effect regression models[5] to the consonant and vowel data. Individual subject and individual phone were included as random effects in the model. Reading disability, phone position and phone type were included as fixed effects. The fixed effects were numerically-coded and centered. All possible interactions between the random effects and fixed effects were examined.

Fixed effect results are listed in Table 3.3. As shown, the regression analysis revealed a similar pattern for fixed effects as the ANOVA. The main effect of

---

[5]Regression analysis was all performed in software R and with the lme4 package.
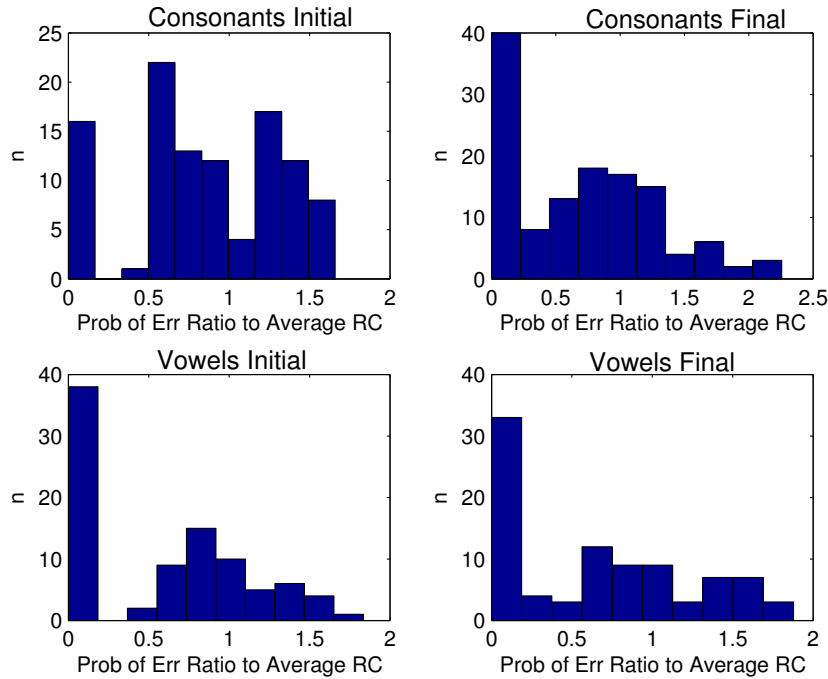
Figure 3.5: Histogram of probability of error ratios of individual RC to average RC for SCO. The general ratios ranged mainly from 0.5 to 1.5

reading disability (b=1.21, SE=0.32, z=3.83, p<.001) was significant, the main effect of phone position (b=-0.19, SE=-.10, z=-1.91, p=.0561) was marginal, and no other fixed effects or interactions between the three main factors are significant. In addition, a couple of random effects were shown to be significant: a main effect of subject (χ^2=746.85, p<.001), interaction between individual subjects and different phone types (χ^2=29.28, p<.001), interaction between individual subjects and different phone positions (χ^2=21.01, p<.001), and interaction between phone type and phone position for different subjects (χ^2=8.01, p<.005).
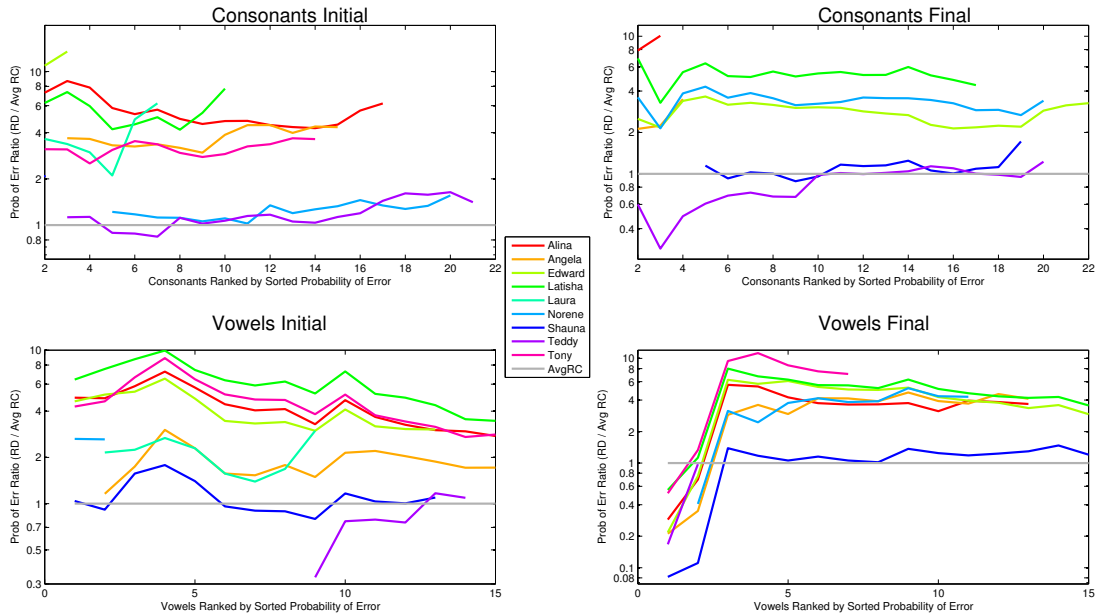
Figure 3.6: Probability of error ratios of individual RD to average RC for SCO. The probabilities of error from individual RD were divided by the probabilities of error from the average RC, and the ratios were plotted in colored curves. A gray line was plotted at 1 for all phones, representing the average RC level. A couple of RDs were performing at the average RC level: RD-Shauna and RD-Teddy in CF; RD-Shauna and RD-Norene in CI; RD-Shauna and RD-Teddy in VI; RD-Shauna in VF. Their ratio curves were separated from other RDs with a clear gap. Some sudden bumps in probabilities of error at the third least error vowel were observed for both initial and final position. The 10th least error vowel at initial position also had a bump-up in ratio value for all RDs. Interestingly, in Fig. 3.4, RCs' ratios all had a dip at the 9th least error vowel.

3.4   Sorted Error Plots for the NSCM Task

3.4.1   Four conditions

RC and RD group patterns of Figure 3.8 are not as well separated for the NSCM task as they are for the SCO task. From comparison to Figure 3.19, we shall see that on average children in the RD group were making more errors than their peers in the RC group under all four conditions (i.e., CI, CF, VI and VF).

Identifying and articulating consonants in both syllable-initial and syllable-final positions comprised a difficult task for all the children. However, syllable-final position proved more difficult, judging from top right panel of Figure 3.8. Thirteen out of 15 children have higher than 10% error when they hit their 10th
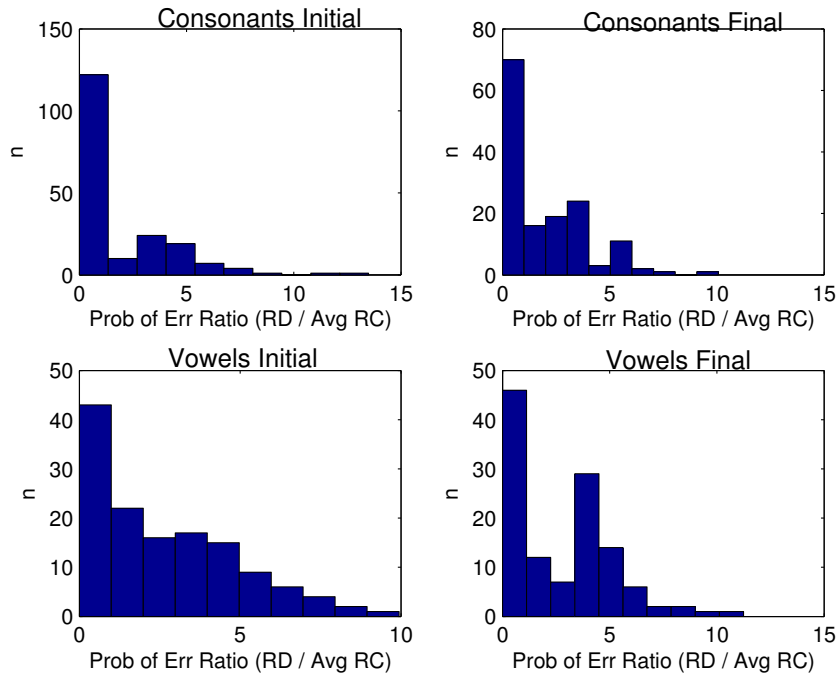
Figure 3.7: Histogram of probability of error ratios of Individual RD to average RC for SCO. The ratios ranged mainly from 1 to 6.

least error consonant in the final position, with the exception of RC-Anton and RC-Evan in the RC group. In the case of identifying and producing consonants in the initial position, all of the children's error rates for their 10th least error consonant were still less than 10%. Out of 24 consonants, the average RC had 15 consonants in the initial position with less than 10% error, and 10 consonants in the final position with less than 10%.

Also, comparing the top left and top right panels of Figure 3.8, the group separation is more obvious for consonants in the syllable-final position. For CF, it appears to be more challenging to recognize even for normal reading children, let alone children with reading disabilities.

The separation in vowel identification and production for the two groups was not big. RD-Laurain the RD stands out for her high error rate in both syllable-initial and syllable-final positions. Very interestingly, however, as demonstrated in later figures for entropy for this task, despite her high error, the entropy of her

Table 3.3: Fixed Effects from Regression Analysis. Generalized Linear Model: Error $\sim$ SubGrp $\times$ PhonePos $\times$ PhoneType.

Fixed Effect	Estimate	Std. Error	z value	Pr($> |z|$)	Sig. Level
(Intercept)	-2.18865	0.16294	-13.433	$<$ 2e-16	***
PhonePos	-0.18853	0.09869	-1.910	0.056084	
SubGrp	1.21389	0.31710	3.828	0.000129	***
PhoneType	-0.02343	0.12473	-0.188	0.851011	
PhonePos:SubGrp	-0.14125	0.18889	-0.748	0.454600	
PhonePos:PhoneType	0.19163	0.17024	1.126	0.260321	
SubGrp:PhoneType	-0.07807	0.21251	-0.367	0.713346	
PhonePos:SubGrp:PhoneType	-0.31775	0.31830	-0.998	0.318151	
Signif. codes	$p < 10^{-3}$	'***';	10^{-2} '**';	0.05 '*'	

error responses for quite a few vowels was relatively low.

3.4.2 Benchmarks generated from the RC group

On average, RCs have 15.44% error for consonants, 22.04% error for vowels, and 18.11% for all the phones. A profile of a hypothetical control child is created from the average performance of the group as presented in Figure 3.9, using a log scale. In other words, the variances of the error rates in perceiving the phones increased exponentially from low error phones to high error phones. Individual RCs performances relative to the performance of this hypothetical average control child are plotted in Figure 3.10. From the histogram of the deviations from average (Figure 3.11), it can be seen that generally the error rate distributions in all four conditions are normal. The ratio of the individual errors to average error ranges from $\approx$0.5 to $\approx$2. The median of the average performance is 0.1216 for CF; 0.0697 for CI; 0.1176 for VF; and is 0.1651 for VI. The standard deviation of the mean for CF is 0.0437; 0.0333 for CI; 0.0508 VF; 0.0617 for VF.
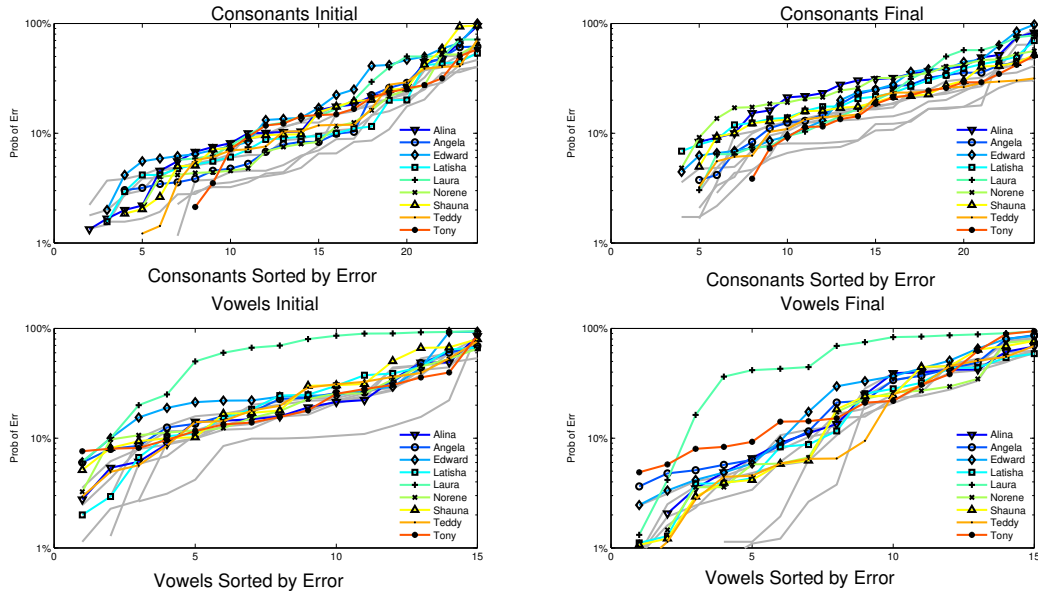
Figure 3.8: Sorted Error Plots for NSCM task. That the curves do not start at 1 on x-axis indicates zero error for those early-ranking phones; and that the curves do not end at 24 (consonants) or 15 (vowels) indicates 100% error for those late-ranking phones. The curves tend to have greater slopes for the consonant initial and vowel final conditions than the consonant final and vowel initial conditions.
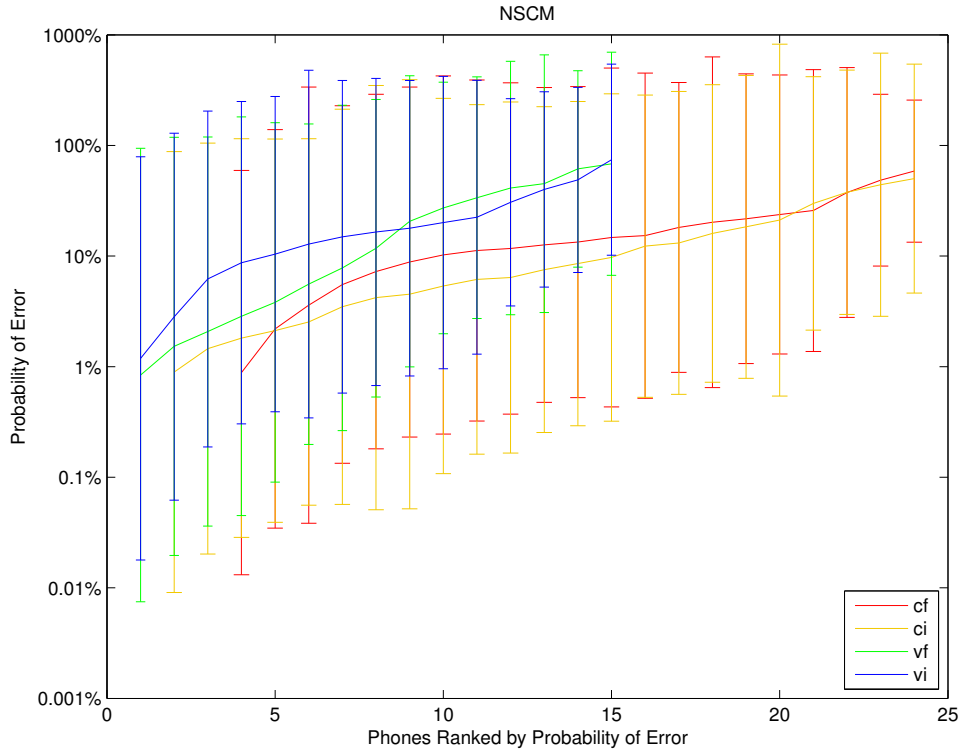
Figure 3.9: Average RC performance in NSCM (log scale). The average RC erred at 18.11% for all phones with stabilized or slightly increased variance.
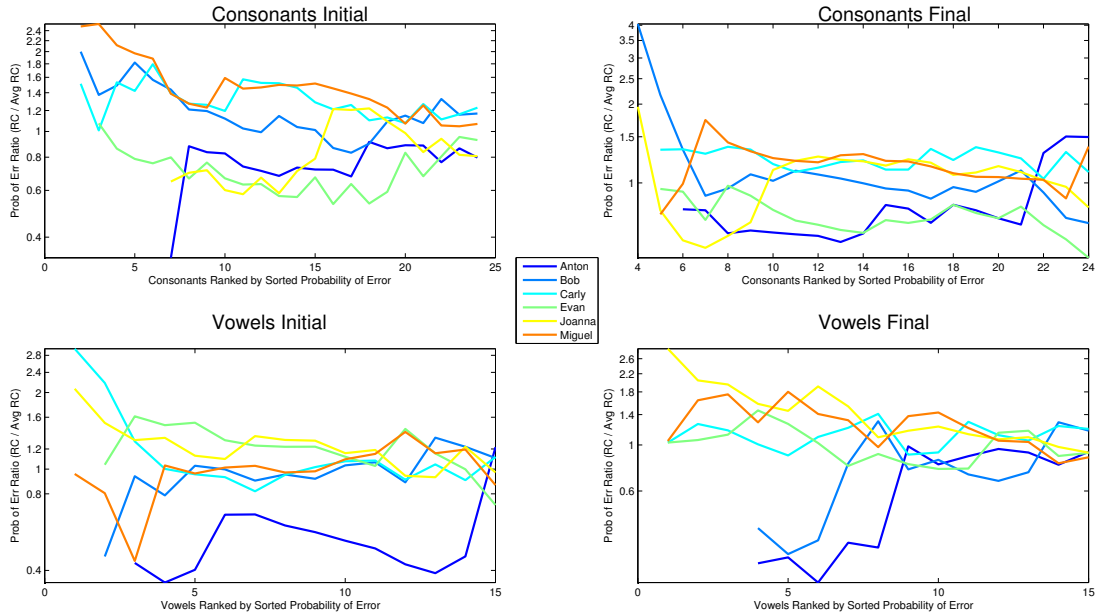
Figure 3.10: Probability of error ratios of individual RC to average RC in NSCM. Among the RC group, Anton was among those with the least overall error. Evan had the least errors for consonants. Almost all curves in each figure converged at certain ranking position: in CI, they started to converge at the sixth least error consonant; in CF, they started at the eighth least error consonant; in VI they started at the fifth least error vowel; in VF they started at the ninth vowel.
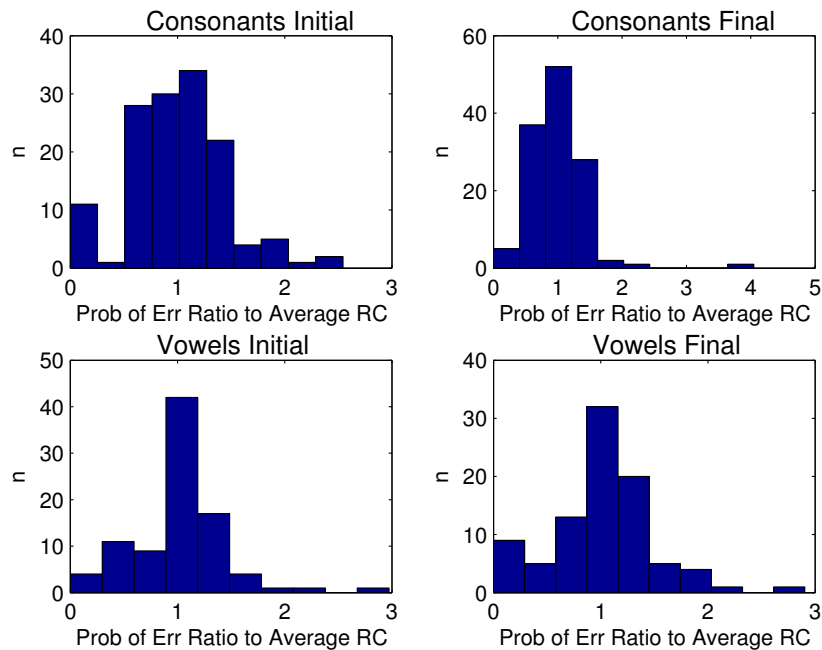
Figure 3.11: Histogram of probability of error ratios of individual RC to average RC in NSCM. Histograms were unimodal for all cases, with value 1 as the most frequent ratio. RC children as a group performed similarly, with few outliers.
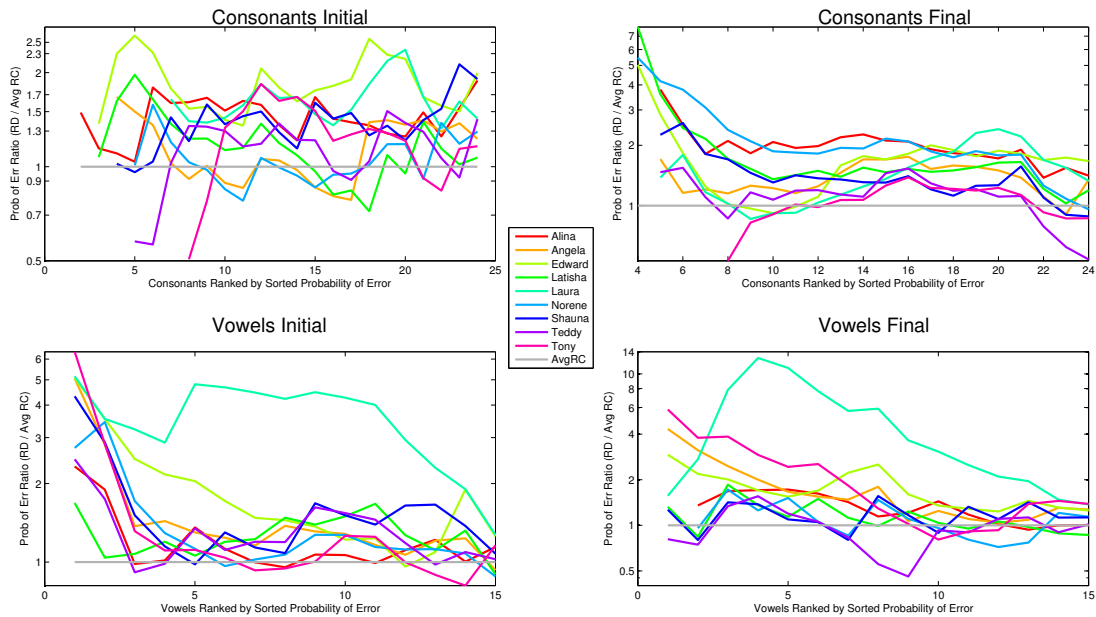
Figure 3.12: Probability of error ratios of individual RD to average RC in NSCM. The probabilities of error from individual RD were divided by the probabilities of error from the average RC, and the ratios were plotted in colored curves. A gray line was plotted at value one for all phones, indicating the average RC level. Mostly, RDs had higher errors than the average RC. Exceptions happened at the first couple of consonants in the initial position: almost every RD erred less than the average RC. In CI, the curves started to converge at the sixth least error consonant; in CF, curves started to converge at the fifth least error consonant; in VI, they started at the third least error vowel; in VF, they started at the fourth least error vowel.

3.4.3 Probabilities of error from RDs and their distributions relative to the average RC level

On average, RDs have 21.27% error for consonants, 29.77% error for vowels, and 24.71% for all the phones. Probabilities of error ratios of RDs to the average RC level and the distribution of the ratios are presented in Figure 3.12 and Figure 3.13. As mentioned in section 3.4.2, inside the RC group, subjects deviate from the mean with ratios ranging from zero to two. The performance of individual RDs ranges mostly from one to three.

3.5 Repeated Measure ANOVA: Result for NSCM

From the result listed in Table 3.4 on page 43,[6] the probabilities of error for the RD children were slightly different from those for the RC children. Error rates for the RD children were greater than for the RC children but only at a significance level of 0.07; error rates for phones in the final position were greater than for phones in the initial position; and error rates for vowels were significantly greater than for consonants. After closer inspection of subject performance, a couple of RDs were found to be performing at the same level as RCs. Therefore a post hoc analysis was performed without the data from these RDs. Results are shown in Table 3.5.

The mean values of all the main effects and significant interactions for Table 3.4 are listed in appendix D.

---

[6]All of the statistic testing in this thesis is performed with R. The probabilities of error were all arcsine transformed before the analysis of variance.

3.6 Entropy Plots from the NSCM Task

3.6.1 Entropy Calculation

Entropies for the error rates in all four conditions were calculated for the NSCM task. Entropy is an indicator of the degree of diversity in responses, or the inconsistency in responses. An intuitive perspective for it is the average odds or average contained information for all the outcomes from the same test, experiment or bet, in the unit of bits. The value of entropy is determined by the amount of information rooted in the source (in our case, the designed experiment) and the processing and delivery capability of the channel and output device (in our case, the speech unit perception capability of the children). So, a high entropy value may reflect that children received too few helpful cues for them to correctly identify the sounds. In other words, if the entropy is high, the child may be guessing about which phone he or she heard. Entropy can be calculated with Equation 3.1. Here, n is the number of total response types.

$$
H = E(log_2 I_{k|j}) \tag{3.1}
$$

$$
= \sum_{k=1}^{n} p_{k|j} log_2 \frac{1}{p_{k|j}} \tag{3.2}
$$

$$
= -\sum_{k=1}^{n} p_{k|j} log_2 p_{k|j} \tag{3.3}
$$

where $I_{k|j}$ is information density as $I_{k|j} = \frac{1}{p_{k|j}}$ and $p_{k|j}$ is the probability of reporting event k given event j (Cover and Thomas, 2006).

In Figure 3.14, red points represent the entropy data versus the error rate data for RD children and green points, represent the same for RC children. The two letters on each point were the short representation for the pseudonym for each child. For example, 'Al' represents RD-Alina; 'At', RC-Anton, etc. (please refer
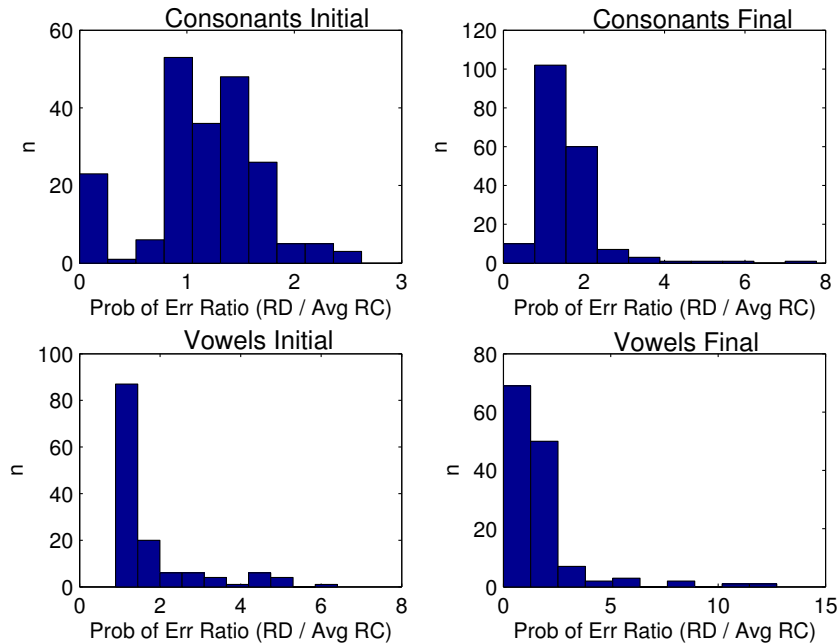
Figure 3.13: Histogram of probabilities of error ratios of individual RD to average RC in NSCM. The values of ratios for all phones mostly concentrated at or slightly higher than value one. That is, RDs were performing at a level close to average RC.

to Appendix C for detailed information). Reference curves of maximum entropy for multiple groups of responses were plotted with solid colored lines. The purple line is for the relation between entropy and the probability of an error response while there are two kinds of response in total: one correct and one wrong. The maximum entropy for two choices is achieved when the two choices have an equal chance of appearance; therefore the curve reaches its maximum point of 1 bit at 50% probability of error for the correct response. The blue line describes the condition for three kinds of responses in total: one correct and two wrong. The maximum entropy in this case would be achieved as 1.58 bits, at two thirds of the probability of error (with one third of the probability being a correct response), where the two confusions share the two thirds equally. The cyan line describes the situation for four kinds of responses: one correct and three wrong. Similarly, it maxes out at 75%. The green curve describes the situation for five kinds of responses: one correct and four wrong. The yellow curve describes the situation

41

for six kinds of responses: one correct and five wrong. The orange curve describes the situation for eight kinds of responses: one correct and seven wrong. The red curve describes the situation for 15 kinds of responses: one correct and 14 wrong. Finally, the brown curve describes the situation for 20 kinds of responses: one correct and 19 wrong. Confusion matrices for the 24 consonants and 15 vowels will be shown in appendix G as stacked bar plots.

The dimensionality of the consonant space was 24. In case of the entropy performance for consonants in syllable-initial position, RC children mostly had 2 to 7 confusions with the correct phone; RD children had 1 to 14. When consonants were placed in the final position of the syllables, alternate confusions for the RC children ranged from 1 to 19 different phones. The same was true for the RD children and they had notably higher error rates. These data demonstrate the difficulty of discerning consonants in syllable-final position compared to syllable initial position, which was consistent with patterns reported by Schuele and Boudreau (2008).

The total dimensionality of vowel space was 15. RCs had 2 to 7 confusions when trying to identify vowels in syllable-initial position. The range was the same as for consonants, however because of the different dimensionality, the confusion diversity for vowels was increased by 12%. For RDs, it was again the same range for the number of confusion types, but with notably higher error rates on more phones than those for RCs. The results for vowels for both RCs and RDs in syllable-final position on the NSCM task were similar to those for vowels in syllable-initial position.

Some preliminary observations can be made from these plots. For extreme cases of entropy, differences exist between consonant perception (the top four plots in Figure 3.14) and vowel perception (the bottom four plots in Figure 3.14). For consonants, we can see that children (both RCs and RDs) were completely lost when the points fell on the curve of 20 alternative responses. This worst case happened when RDs identified consonants in syllable-final positions. They gave

Table 3.4: Repeated Measure ANOVA Result for NSCM task. Probabilities were processed by arcsine transform. Linear Model: ProbOfErr ~ SubGrp * PhonePos * PhoneType.

Factor	Df	Sum Sq	Mean Sq	F value	Pr($>$F)	Sig. Level
(Individual)						
SubjectGroup	1	17297	17297	8.025	0.0141	*
Residuals	13	28020	2155			
(Within Individual)						
PhonePosition	1	3235	3235	6.549	0.01062	*
PhoneType	1	14772	14772	29.903	5.57e-08	***
SubjectGroup:PhonePosition	1	1	1	0.002	0.96149	
SubjectGroup:PhoneType	1	0.00	0.0000	0.000	0.99420	
PhonePosition:PhoneType	1	5217	5217	10.561	0.00119	**
sub:pos:type	1	188	188	0.381	0.53745	
Residuals	1149	567606	494			
Signif. codes	$p < 10^{-3}$	'***';	10^{-2}	'**';	0.05	'*'

many alternative choices and performed at almost chance level. However, this did not happen for the vowels. In total, there were 15 vowels, and the worst vowel cases even for RDs only reached 8 alternatives for 1 vowel. This happened for vowels in syllable-initial position. The point to be made here is that there are consonants that can be so confusing that in certain syllable positions, children hardly recognize them. However, vowel perception maintains certain boundaries that normal listener seldom cross, so that normal listeners seldom make completely random guesses for vowels. Another observation is that in the low entropy region for the RD group, quite a few children, like RD-Laura , had some phones that were highly confusable with only one other phone. In the RC group, phones with only one confusion were present, but only with less than a 20% probability of error.

3.6.2 Closer Look at Entropy

Figure 3.15 is a display of phonetic perception for phones with only one confusion. In all conditions, the error rates for phones with only one confusion were mostly below 30%.

Figure 3.16 is a display of phonetic perception for phones with about two con-

Table 3.5: Post Hoc Repeated Measure ANOVA Result for NSCM task: RD-Teddy and RD-Norene were deleted from CI; RD-Shauna and RD-Teddy were deleted from CF; RD-Alina and RD-Teddy were deleted from VI; RD-Norene and RD-Teddy were deleted from VF. Probabilities were processed by arcsine transform. Linear Model: ProbOfErr ~ SubGrp * PhonePos * PhoneType.

Factor	Df	Sum Sq	Mean Sq	F value	Pr(>F)	Sig. Level
(Individual)						
SubjectGroup	1	22378	22378	9.698	0.011	*
PhonePosition	1	126	126	0.054	0.820	
PhoneType	1	123	123	0.053	0.822	
Residuals	10	23075	2308			
(Within Individual)						
PhonePosition	1	3245	3245	6.382	0.01168	*
PhoneType	1	14445	14445	28.405	1.22e-7	***
SubjectGroup:PhonePosition	1	15	15	0.029	0.86426	
SubjectGroup:PhoneType	1	50	50	0.097	0.75504	
PhonePosition:PhoneType	1	4248	4248	8.353	0.00393	**
sub:pos:type	1	208	208	0.409	0.52287	
Residuals	994	505503	509			
Signif. codes	$p < 10^{-3}$	'***';	10^{-2}	'**';	0.05	'*'

fusions. These phones were located in the region between the curve of exactly one confusion and the curve of exact two confusions.

Figure 3.17 is a display of phonetic perception for phones with many three-way confusions. Both groups had three confusions for the CI, CF, VI, and VF conditions. The children who had phones with high error rates and three confusions were RC-Evan, RC-Carly, RC-Bob, and RC-Miguel; and RD-Tony, RD-Laura, RD-Edward, RD-Shauna, RD-Norene, RD-Alina, and RD-Teddy (i.e., everyone except RD-Angela and RD-Latisha).
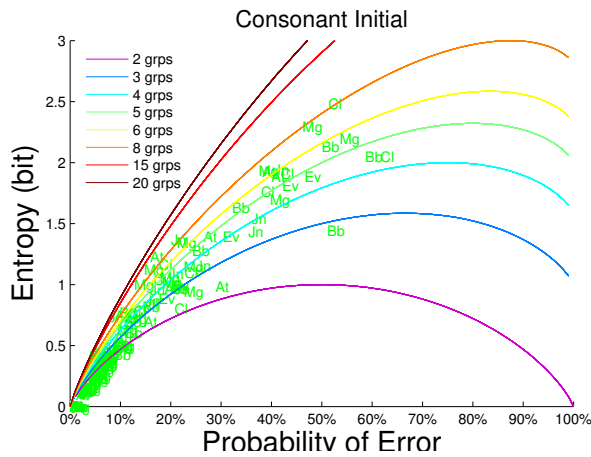
(a) The performance of RCs in the Consonant Initial condition are plotted in green dots. The number of confusions ranges from one to seven.
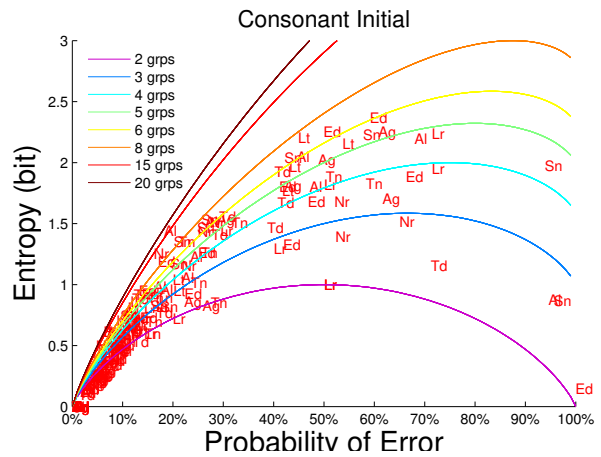
(b) The performance of RDs in the Consonant Initial condition are plotted in red dots. The number of confusions ranges from one to seven.
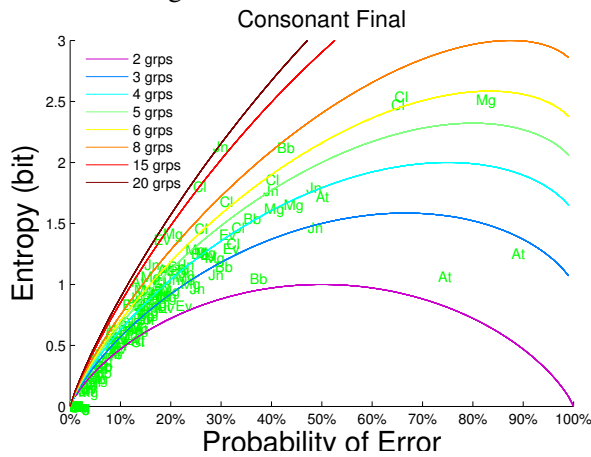
(c) The performance of RCs in the Consonant Final condition are plotted in green dots. The number of confusions ranges from one to 19.
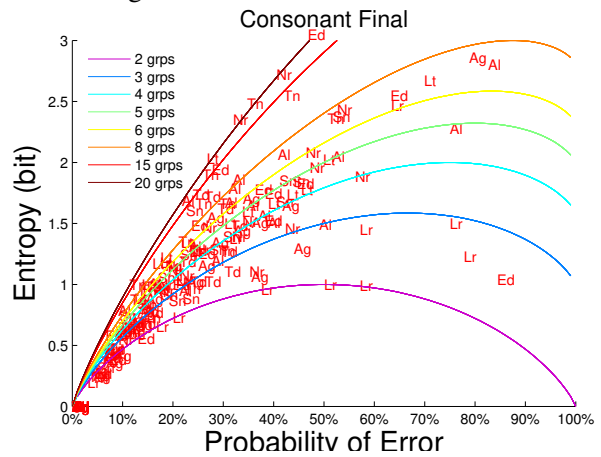
(d) The performance of RDs in the Consonant Final condition are plotted in red dots. The number of confusions ranges from one to 19.

Figure 3.14: Entropy of the probabilities of error for NSCM task.
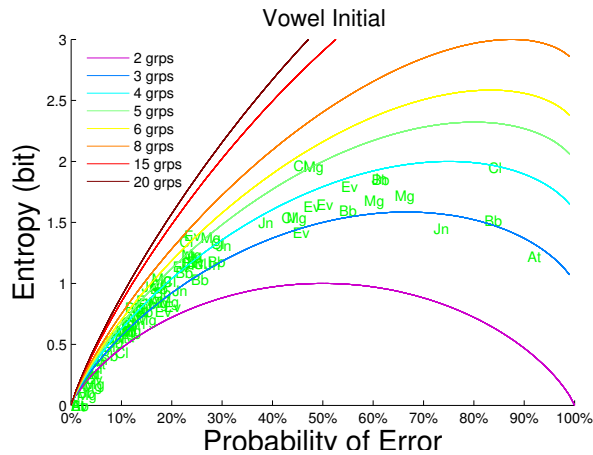
(e) The performance of RCs in the Vowel Initial condition are plotted in green dots. The number of confusions ranges from two to about four.
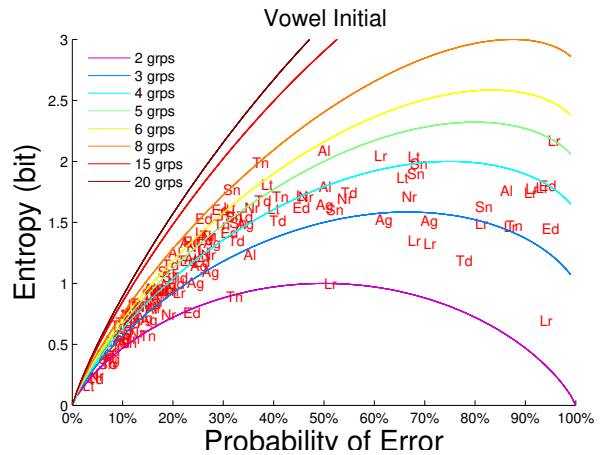
(f) The performance of RDs in the Vowel Initial condition are plotted in red dots. The number of confusions ranges from more than one to seven.
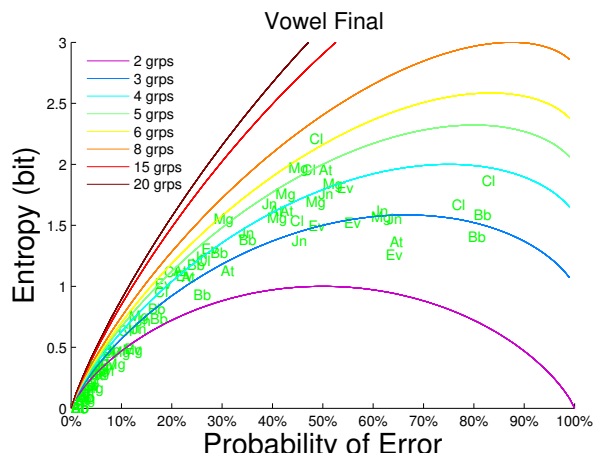
(g) The performance of RCs in the Vowel Final condition are plotted in green dots. The number of confusions ranges from one to five. Probabilities of error are generally below 60%, with four exceptions at about 80%.
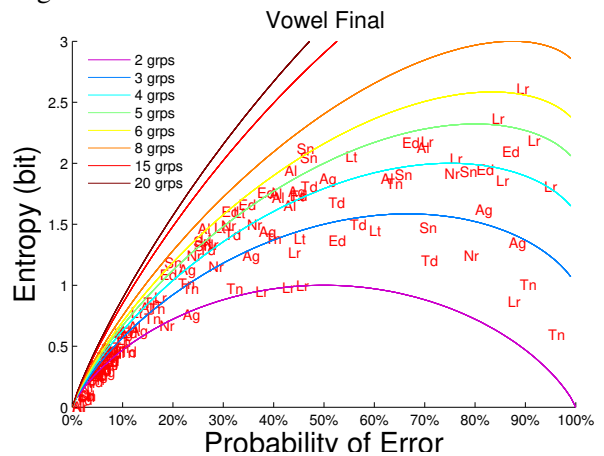
(h) The performance of RDs in the Vowel Final condition are plotted in red dots. The number of confusions for all the vowels in the final position ranges from one to six. There are more high error vowels for RDs than RCs. Probabilities of error range from zero to more than 90%.

Figure 3.14: cont.

46
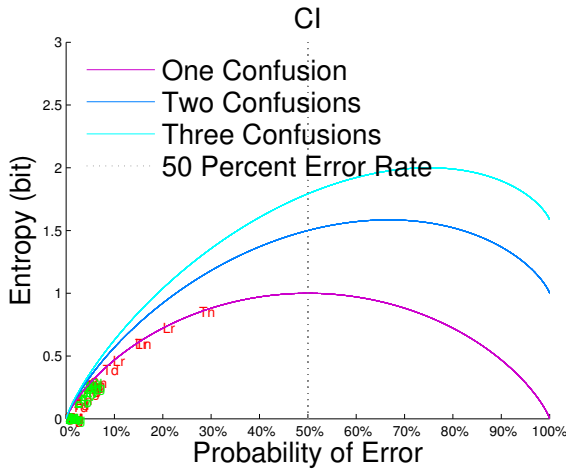
(a) One Confusion Summary for CI: most one-confusion phones for the RCs have less than 10% error. RDs who have slightly high error one-confusion phones are Tony, Laura, and Teddy.
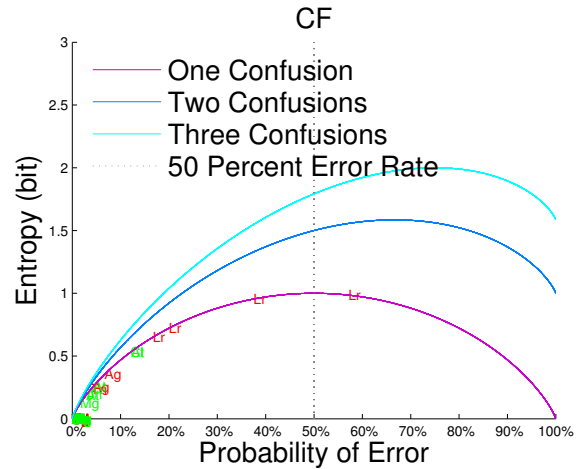
(b) One Confusion Summary for CF: most one-confusion phones for the RCs have less than 10% error. RDs who have slightly high error one-confusion phones are Laura and Angela.
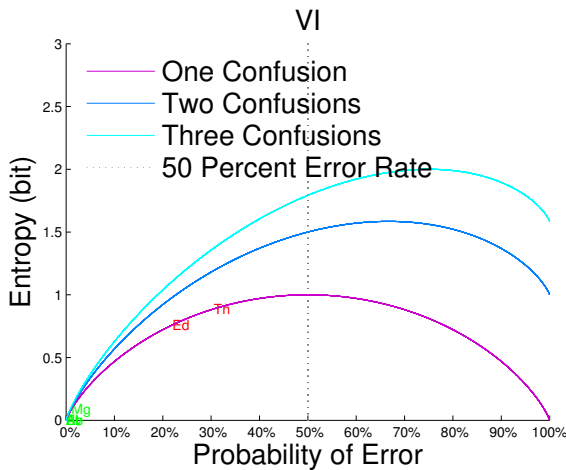
(c) One Confusion Summary for VI: most one-confusion phones for the RCs have 0% error. RDs who have slightly high error one-confusion phones are Edward and Teddy.
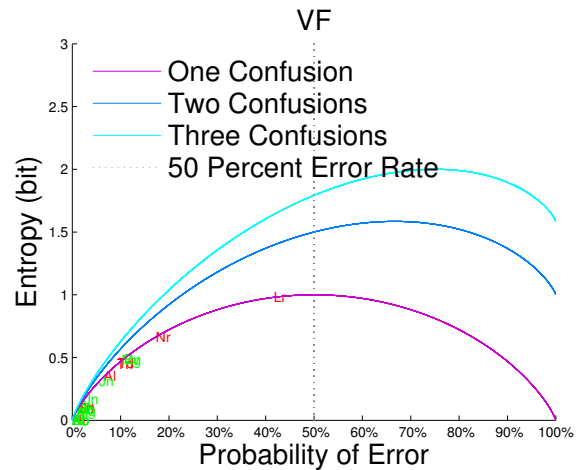
(d) One Confusion Summary for VF: most one-confusion phones for the RCs have less than 10% error. RDs who have slightly high error one-confusion phones are Norene, Laura, and Teddy.

Figure 3.15: One Confusion Summary for the NSCM task.
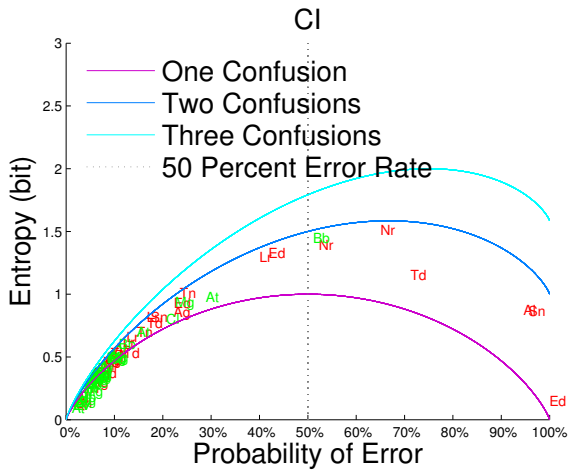
(a) Two Confusion Summary for CI: most two-confusion phones for the RCs have less than 30% error. RDs who have slightly high error two-confusion phones are Norene, Edward, Tony, Laura, and Teddy.
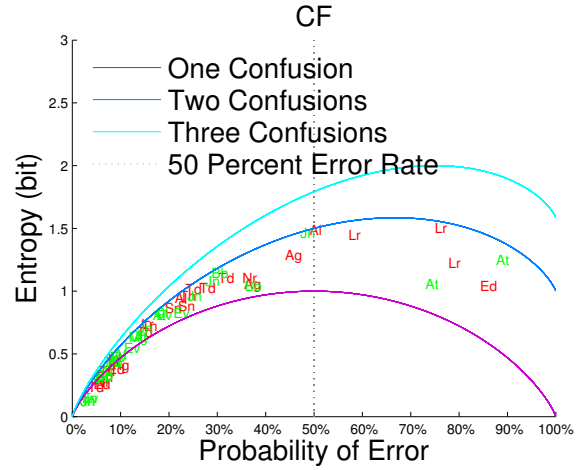
(b) Two Confusion Summary for CF: most two-confusion phones for the RCs have less than 50% error. RDs who have slightly high error two-confusion phones are Alina, Norene, Edward, Angela, Laura, and Teddy.
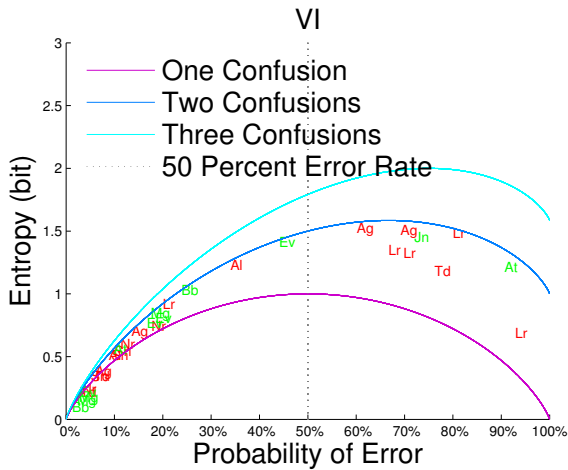
(c) Two Confusion Summary for VI: most two-confusion phones for the RCs have less than 50% error. RDs who have slightly high error two-confusion phones are Alina, Angela, Laura, and Teddy.
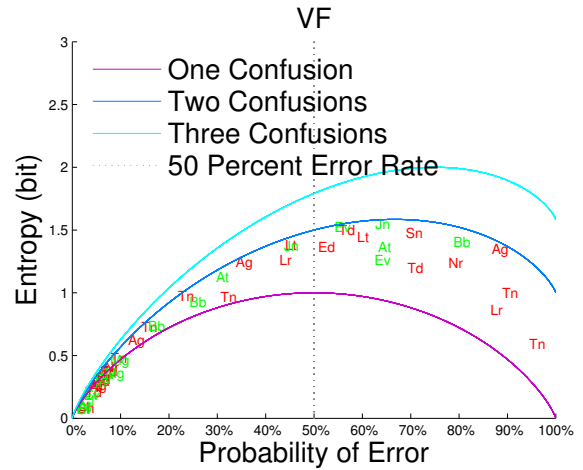
(d) Two Confusion Summary for VF: most two-confusion phones for the RCs have less than 70% error. RDs who have slightly high error two-confusion phones are Tony, Latisha, Norene, Edward, Angela, Laura, and Teddy.

Figure 3.16: Two Confusion Summary for the NSCM task.
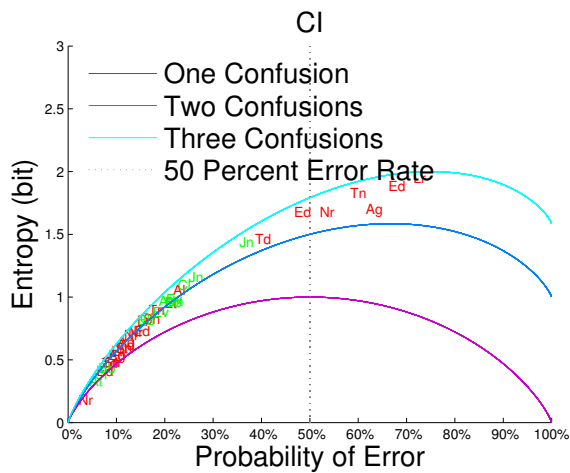
(a) Three Confusion Summary for CI: all three-confusion phones for the RCs have less than 40% error. RDs who have slightly high error three-confusion phones are Tony, Norene, Edward, Angela, Laura, and Teddy.
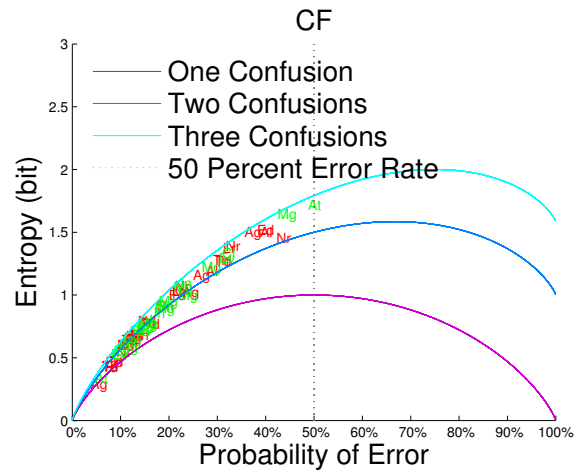
(b) Three Confusion Summary for CF: all three-confusion phones for the RCs and the RDs have less than 50% error. RDs and RCs have performance overlap in this condition. Anton and Miguel in RC and Norene Edward and Angela in RD have slightly high error in this condition.
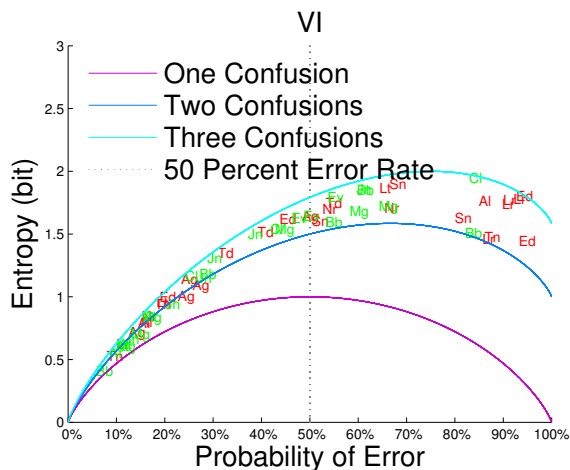
(c) Three Confusion Summary for VI: both the RDs and the RCs have high error phones and high confusable phones in this condition. Children who have slightly high error three-confusion phones are Edwards, Laura, Shauna, Carly, Bob, Norene, and Miguel.
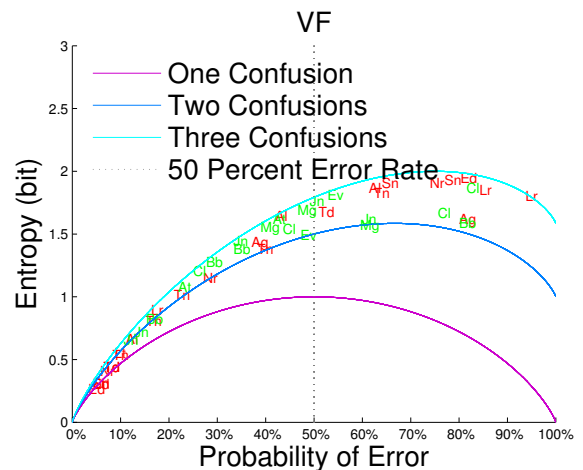
(d) Three Confusion Summary for VF: both the RDs and the RCs have high error phones and high confusable phones in this condition. Children who have slightly high error three-confusion phones are Edwards, Laura, Shauna, Carly, Bob, Norene, Alina, Tony, Teddy, and Miguel.

Figure 3.17: Three Confusion Summary for the NSCM task.

3.7 Correlation between Reading Ability and Phonetic Perception Tasks

After the inspection of the raw results from the phonetic perception tasks, it is important to link these results to the reading abilities of the two groups of children.

As mentioned in the last chapter, reading abilities of all children were assessed by a battery of reading tests. It is necessary to know if the results of the two phonetic perception tasks have any relation to the children's reading levels. Results reported in this section came from previous analysis by Johnson *et al.* (2011a,b). These tests were the Woodcock Reading Mastery Tests-Revised (WRMT-R), which included the Word Identification subtest (WI) and Word Attack subtest (WA); and the Gray Oral Reading Test, 4th edition (GORT-4), which included a Fluency score (R-FLU) and a Comprehension score (R-COMP). Word recognition abilities were measured by the scores on the WI and WA subtests, and other higher level reading skills were by the R-FLU and R-COMP scores.

The scores from SCO and NSCM tasks were compared with the scores from the four reading tests to calculate the correlation between phonetic perception and reading ability. Pearson correlation coefficients are displayed in Table 3.6. The number of stars following the coefficients indicated the significance level of the correlation between two measures. The results from the SCO task had correlated significantly with the results of all four reading tasks but the results from the NSCM task did not, showing again that the SCO task is more effective than the NSCM task in separating the RC from the RD group. The fact that the scores from the SCO task correlated more highly with word recognition and reading fluency than reading comprehension also demonstrated a closer linkage of phonetic perception with lower level components required in normal reading ability than the linkage with higher level cognitive components in reading ability.

Table 3.6: Correlation between Phonetic Perception and Reading Tests

	WI	WA	R-FLU	R-COMP
SCO	0.778(**)	0.729(**)	0.773(**)	0.617(*)
NSCM	0.314	0.442	0.392	0.379
Signif. codes	$p < 10^{-3}$ '***';	$p < 10^{-2}$ '**';	$p < 0.05$ '*'	

3.8 Confusion Examples from Subjects on the NSCM Task

With a general conclusion of significant group difference in the two tasks, it is important to also know the detailed perceptual patterns from the two groups. The NSCM task provides these detailed perceptual confusion patterns.

Figure 3.19a consists of the phonetic perception of syllable-initial consonants for four children: the child with the best in the RC group (RC-Evan), the best in the RD group (RD-Norene), the worst in the RC group (RC-Miguel) and the worst performance in the RD group (RD-Edward). The stars and circles were the average performances generated by the two groups, respectively. With the same display style, Figure 3.19b shows the best and the worst performances for the two groups for syllable-final consonants, 3.19c for syllable-initial vowels and 3.19d for syllable-final vowels. Confusion matrices are shown as stacked bar plots for individuals in Fig. G.1, G.2, G.3, G.4, G.5, G.6, G.7, and G.8. The confusion matrices were sorted along the x-axis by probability of erroneous identification of the target phones. The target phone with the least error sits farthest left; the target phone with the most errors sits farthest right of the axis. Because each child had a different degree of correct perception for each phone, the sequence of phones arranged along the x axis was different for each child. The benefit of using this type of visualization is that it offers a direct view of which phones require more of the child's attention and, therefore, where to focus time and intervention for the individual child with a reading disability.

Another way to visualize the confusion patterns is to treat perception as a dynamic flow between the whole set of phones, as in Figure 3.18. That means to view confusing alternative phones as possible transition paths from the presented

phone. Following this direction, the confusion matrices become stochastic transition matrices and describe perceptual state transitions. Therefore, stochastic transition matrices for each child were plotted as directed diagrams and are presented in Appendix E [7].

3.8.1 Average RC Confusion Patterns

To understand the average phonetic perception for normal reading children, the mean values of confusion percentages between phones were extracted from the RC data (see Figure 3.20a, Figure 3.21a, Figure 3.22a, and Figure 3.23a on page 56, 58, 60, and 61). From these confusion patterns, we can understand the obstacles that exist in perceiving natural English for normal reading children. Analyzing difficult points in phonetic perception for a specific language in its spoken form could help us understand the perceptual limits of normal readers from normal language and educational backgrounds. It is interesting and important to see that even with a certain degree of deviation in perception, the speech in this kind of language is sufficiently robust at the bottom layer of speech processing for us to communicate.

AVERAGE RC CONFUSION PATTERNS FOR CI

As shown in Figure 3.20a, children with normal reading abilities could identify all presented phones with at least 54% accuracy, and had at least 91% accuracy in identifying phones such as /g/, /h/, /k/, /l/, /m/, /n/, /p/, /r/, /t/, /w/, /s/, /ʃ/, /tʃ/, /d/, /ʤ/, and /j/ when these consonants were put in the syllable-initial position. Confusion patterns that appeared more than 10% of the time were generally for fricative and affricate targets: θ→ð (16%), θ→f (16%), ʒ→ʤ(16%), ð→θ (14%), f→θ (11%), v→ð (11%). Confusion patterns that appeared less than 10% of the time

---

[7]The open source program 'GraphViz' used in generating these graphs is from AT&T research.
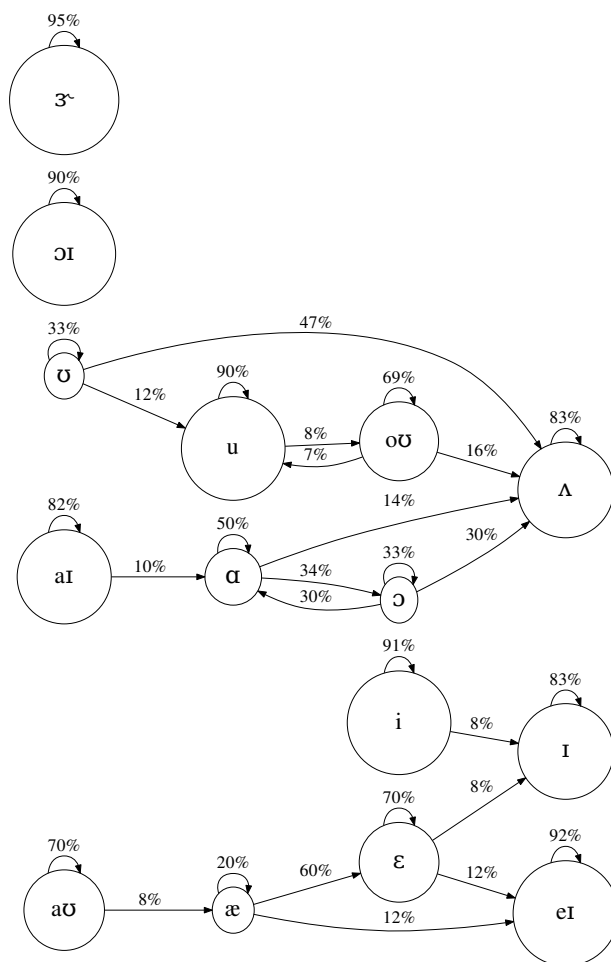
Figure 3.18: Example of a directed graph: Shauna (RD) VI. Each arrow represents a transition patterns, and the percentage accompanying the arrow represents the probability for that transition. Arrows which go back to their original circles represent a correct perception for the phones. The size of circle is proportional to the probability of correct perception of the phone inside of the circle. Therefore, the larger number accompanying the self-loop, the larger the circle is, meaning larger percentage of correct perception of this phone.
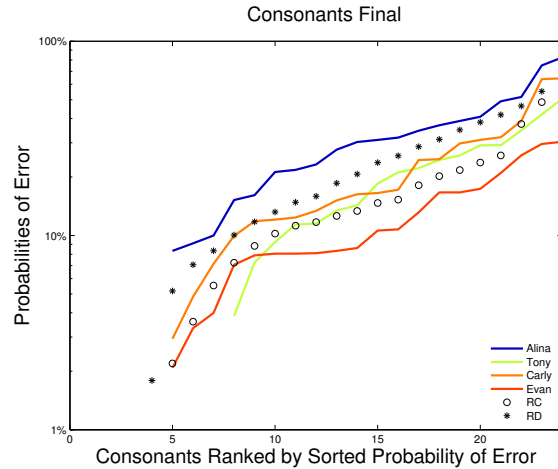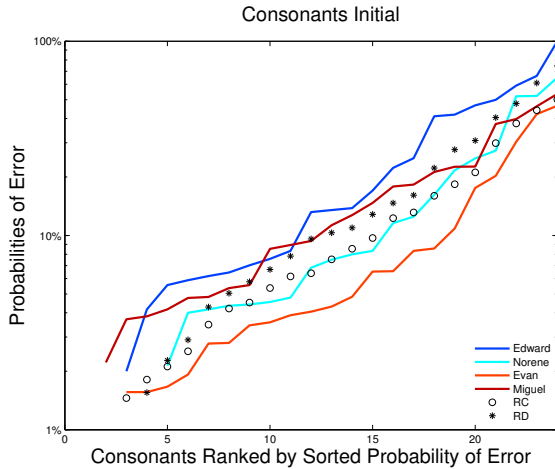
(a) Consonant Initial: Best RC: Evan; Best RD: Norene; Worst RC: Miguel; Worst RD: Edward.

(b) Consonant Final: Best RC: Evan; Best RD: Tony; Worst RC: Carly; Worst RD: Alina.
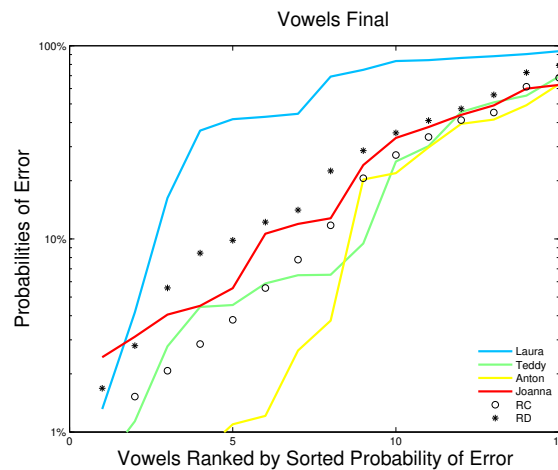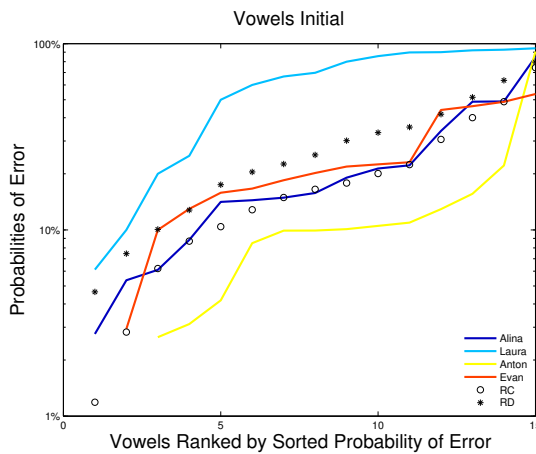
(c) Vowel Initial: Best RC: Anton; Best RD: Alina; Worst RC: Evan; Worst RD: Laura.

(d) Vowel Final: Best RC: Anton; Best RD: Teddy; Worst RC: Joanna; Worst RD: Laura.

Figure 3.19: Selected children from NSCM: Blue and green colored lines are selected RD children performance for all four cases; Red, yellow and orange colored lines are selected RC children performance for all four cases; Black stars are average performance for RD group; Black circles are average performance for RC group. As presented, on average, the RC group performs better than the RD group. However, in each condition, the best RD performed at the same level as the worst RC, if not better.

54

but were still notable happened between obstruents (sibilants, fricatives, affricates, and stops): v→f (9%), z→s (9%), ʒ→ʃ (9%), θ→s (8%), ð→v(7%), ð→d(6%), v→b (6%), ʤ→tʃ (5%), and z→ʒ (5%).

Based on the connections depicted in Figure 3.20a, there was a clear separation between these confusing phones and those intact phones like nasals, glides and stops; however, inside the confusion patterns, there were no obvious boundaries for phones with different features. They all belong to a large group. Phones at the center of confusions were /θ, ð, f, v/ and /ʒ/.

AVERAGE RC CONFUSION PATTERNS FOR CF

In Figure 3.21a, on average for consonants in final position, children with normal reading abilities could identify all presented phones with at least 54% accuracy; and had at least 81% accuracy for intact phones (phones that have no or very few confusion arrows coming out of them) such as /ʃ/, /tʃ/, /b/, /k/, /t/, /l/, /p/, and /r/. Compared with consonants in initial position, there were fewer intact phones in final position. Smaller circles around the phones indicate smaller correct recognition for them in normal reading children. These weaker phones were /ð/, /ŋ/, /θ/, /f/, /v/, /ʒ/, /n/ when they were in final position.

Confusion patterns that happened more than 20% of the time emerged in syllable-final position: f→θ (23%) and ŋ→n (20%). Less severe confusion patterns that happened more than 10% of the time were: ð→v (17%), v→ð (17%), θ→f (12%), ʒ→ʃ (11%), and m→n (10%). Confusion patterns that happened less than 10% of the time but were still notable were: n→m (9%), ð→θ (9%), s→θ (9%), θ→s (8%), z→s (7%), b→v (6%), n→ŋ(6%), ŋ→m (6%), g→k (6%), ʒ→ʤ(5%), ʤ→tʃ (5%), ŋ→g (4%), and d→t (4%).

Unlike the previous condition where consonants were put in initial position, the confusing phones in final position were mixed with a few intact phones. Also, another significant observation about boundaries of perception could be made: A
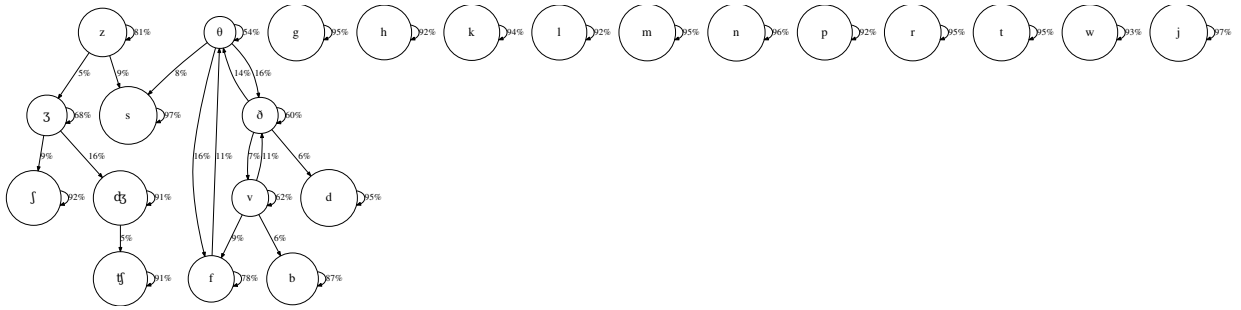
(a) Average RC Phonetic Confusion Patterns for Consonant at Initial Position Condition. There was a clear separation between these confusing phones and those intact phones like nasals, glides and stops, however, inside the confusion patterns, there were no obvious boundaries for phones with different features. Phones in the center of confusions were θ, ð, f, v, and ʒ. Intact consonants were g, h, k, l, m, n, p, r, t, w, s, ʃ, tʃ, d, ʤ, and j, each with at least 91% correct identification rate.
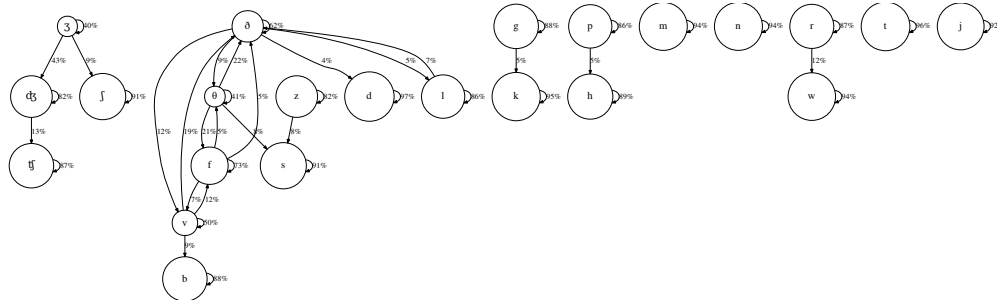
(b) Average RD Phonetic Confusion Patterns for Consonant at Initial Position Condition. The confusion patterns and the categorization between confusing phones were very similar to RC performance, only with bigger confusing rates. One category was formed by phones: ʒ, ʤ, ʃ, and tʃ. All four were palatal obstruents. Another category consisted phones: ð, θ, f, v, b, s, z, d and l. Except for l, all of these were fricatives, sibilants and stops. Mostly the confusions were concentrated in fricatives, affricates, sibilants and two stops. Other very different confusion patterns that were not observed in average RC were ð→l, l→ð, g→k, r→w and p→h; while z→ʒ in RC group did not happen here. Since these six patterns had only less than 10% transition rate, the main structure of all RD confusion patterns remained similar to that of RC's. Confusing centers were θ, ð, f, and v. Intact consonants were tʃ, ʃ, b, s, d, h, m, n, w, t, and j, each with at least 87% correct rate.

Figure 3.20: Average RC and RD Confusion Patterns in CI

clear separation between categories of phones emerged. The first category consisted of consonants that required more effort or tension on the tongue blade: /ʒ, ʤ, ʃ/ and /tʃ/. The second category included phones such as /ð, θ, f, b, v, s/ and /z/, which were phones that either require participants to use the upper teeth or phones that are sibilants. Nasal and velar phones /ŋ, g, k, m/ and /n/ made the third category of confusing phones. However, these phones were generally intact when presented in syllable-initial position. The fourth category of sounds consisted of /d/ and /t/ and the data show that even for less frequently made confusions, normal reading children sometimes perceived /d/ as /t/ when /d/ was put at the end of the syllable, but not the other way around. Phones at the center of confusions were /ð, θ, s, v, n, m/ and /ŋ/. Related information about articulatory gestures can be found in Shriberg and Kent (1982).

AVERAGE RC CONFUSION PATTERNS FOR VI

Benchmark patterns for vowels in initial position are presented in Figure 3.22a. We can see that for RC children, intact vowels could be correctly recognized at least 82% of the time. These vowels were /ʌ/, /ɔɪ/, /ɝ/, /aʊ/, /aɪ/, /eɪ/, /oʊ/, and /u/. The most difficult vowels were /ɑ/ (only 31% correct), /ʊ/ (50%), /ɔ/ (62%), /æ/ (69%), and /ɛ/ (76%).

For the confusion patterns, one stood out among all: ɑ→ɔ (49%). The vowels sounded so alike that half of the time the RC children would perceive /ɑ/ as /ɔ/. The reverse confusion pattern ɔ→ɑ (19%) was also substantial for the RC children, but less severe than the other way around. About a third of the time /ʊ/ would be perceived as /ʌ/: ʊ→ʌ (34%). Other confusion patterns were: æ→ɛ (21%), ɑ→ʌ (15%), ɔ→ʌ (13%), ɪ→i (9%), ɛ→ɪ (7%), ɪ→ɛ (7%), and i→ɪ (7%).

In sum, when vowels were presented at initial position, the confusing phones formed two groups and they were clearly separated from the intact vowels, except for ʌ. There was a boundary between group one (/æ, ɛ, ɪ, i/) and group two (/ʊ, ʌ,
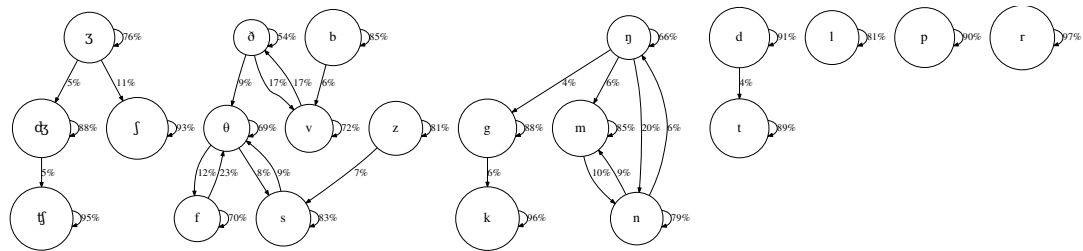
(a) Average RC Phonetic Confusion Patterns for Consonant at Final Position Condition. The first confusion category was consisted of palatal obstruents: /ʒ, ʤ, ʃ/ and /tʃ/. The second category included phones such as /ð, θ, f, b, v, s/ and /z/, which were anterior obstruents. Nasal and velar phones /ŋ, g, k, m/ and /n/ made the third category. However, these phones were generally intact when presented at syllable initial position. The fourth category consisted d and t and the data showed that even less frequently made, normal reading children could perceive d as t when d was put at the end of the syllable but not the other way around. Phones at the center of confusions were /ð, θ, s, v, n, m/ and /ŋ/. Intact consonants were /ʃ, tʃ, b, k, t, l, p/ and /r/, each with at least 81% correct identification rate.
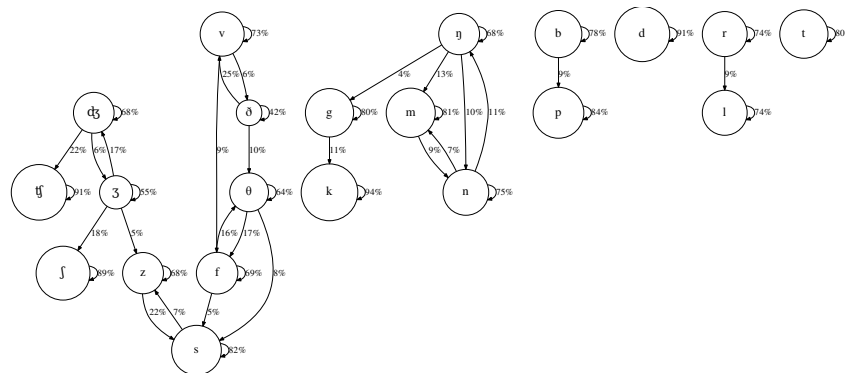
(b) Average RD Phonetic Confusion Patterns for Consonant at Final Position Condition. Again, the confusion patterns were similar to those of RCs'. Some very different confusion patterns that were not observed in average RC confusion patterns were ʒ→z, s→z, b→p and r→l; while patterns like d→t, b→v in RC group did not happen here. Confusing centers were /ʤ, ʒ, z, s, f, θ, ð, v, ŋ, n/ and /m/. Intact consonants were /t, l, d, p, k, tʃ/ and /ʃ/, each with at least 74% correct rate.

Figure 3.21: Average RC and RD Confusion Patterns in CF

ɑ, ɔ/). All phones from group one are front vowels and all phones from group two are all back vowels. Phones at the center of confusions were /ɛ, ɪ, ʌ, ɑ/ and /ɔ/.

In Figure 3.23a, children with normal reading abilities had at least 92 to 98% correct identification of phones such as /ɔɪ/, /ɝ/, /aʊ/, /aɪ/, /eɪ/, /i/, /oʊ/, and /u/ when these vowels were put in the syllable-final position. Similar to the case for syllable-initial position, the most severe confusion patterns were ʊ→ʌ (49%) and ɑ→ɔ (48%). Other confusion patterns were: æ→ɛ (27%), ɔ→ɑ (21%), ɪ→ɛ (19%), ɑ→ʌ (14%), ɛ→æ(11%), ʌ→ɑ (9%), ɛ→ɪ (8%), and æ→ɑ (7%).

Confusing phones did not separate into groups in syllable-final case. Intact phones were all diphthongs and tense vowels when vowels were at final position. This was the same as for vowels initial position. Phones at the center of confusions were /æ, ɛ, ʌ, ɑ/ and /ɔ/.

3.8.2   Average RD Confusion Patterns

To understand the common confusion patterns for RD children, directed graphs were generated for average percent of phonetic perception in four conditions. These are displayed in Figure 3.20b, 3.21b, 3.22b, and 3.23b on page 56, 58, 60, and 61.

As displayed in Figure 3.20b, for the RD children, the intact consonants in syllable-initial position could be perceived with at least 87% accuracy. These phones were /tʃ/, /ʃ/, /b/, /s/, /d/, /h/, /m/, /n/, /w/, /t/, /k/, and /j/. Difficult phones were /ʒ/ (only 40% correct identification, on average), /θ/ (41%), /v/ (50%), and /f/ (73%).
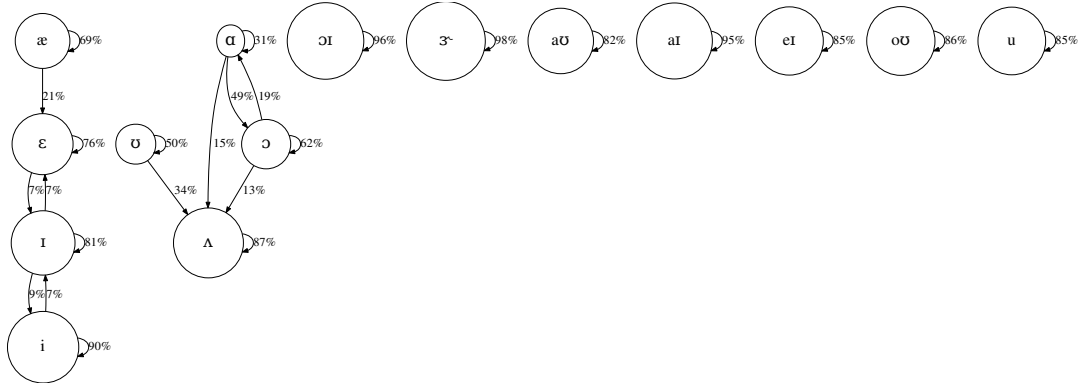
(a) Average RC Phonetic Confusion Patterns for Vowel at Initial Position Condition. The first confusion category consisted lax 'front' vowels /æ, ɛ, ɪ,i/ and the second confusion category comprised of lax 'back' vowels/ʊ, ʌ, ɑ, ɔ/. Confusing centers were /ɛ, ɪ, ʌ, ɑ/, and /ɔ/. Intact vowels were diphthongs and tense vowels such as /ʌ, ɔɪ, ɝ, aʊ, aɪ, eɪ, oʊ/, and /u/, each with at least 82% correct identification rate.
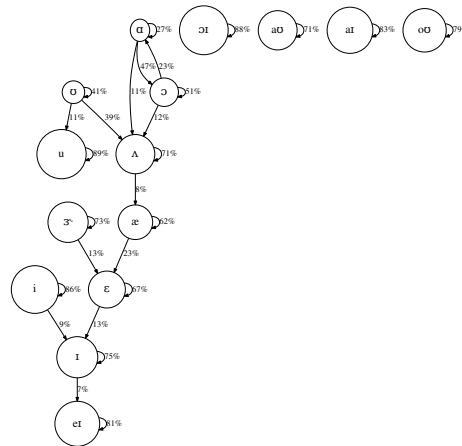
(b) Average RD Phonetic Confusion Patterns for Vowel at Initial Position Condition. Among the four cases (CI, CF, VI, and VF), the general structure of the average RD confusion patterns for VI differed the most from the structure of the average RC confusion patterns. There were six intact vowels for RDs: /u, eɪ, ɔɪ, aʊ, aɪ/, and /oʊ/, with correct rate of at least 71%. However, there was a huge chain of confusion patterns that involved 11 vowels. Phones at the center of confusions were /ʌ, ɛ/, and /ɪ/. Diphthongs like /ɔɪ, aʊ/, and /ɔɪ/ were still robust and did not have significant confusion patterns to other phones. No boundary was found to form groups like those for the RC confusion patterns.

Figure 3.22: Average RC and RD Confusion Patterns in VI
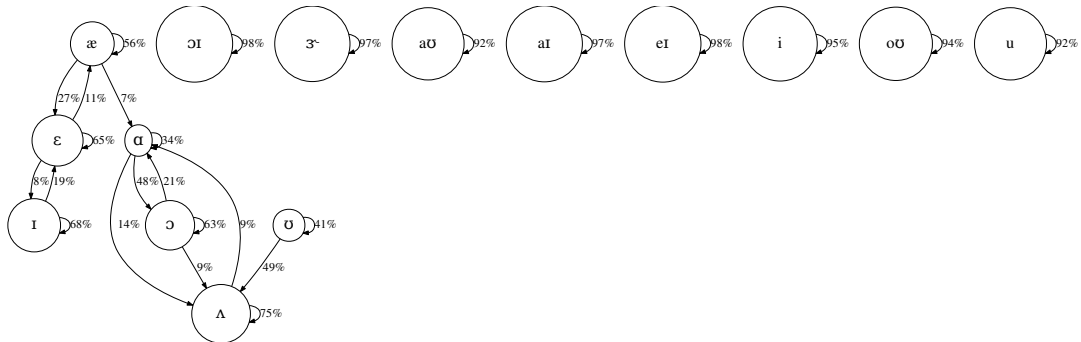
(a) Average RC Phonetic Confusion Patterns for Vowel at Final Position Condition. Confusing centers were /æ, ɛ, ʌ, ɑ/, and /ɔ/. Intact vowels were diphthongs and tense vowels such as /ɔɪ, ɝ, aʊ, aɪ, eɪ, i, oʊ/, and /u/, each with at least 92% correct identification rate.
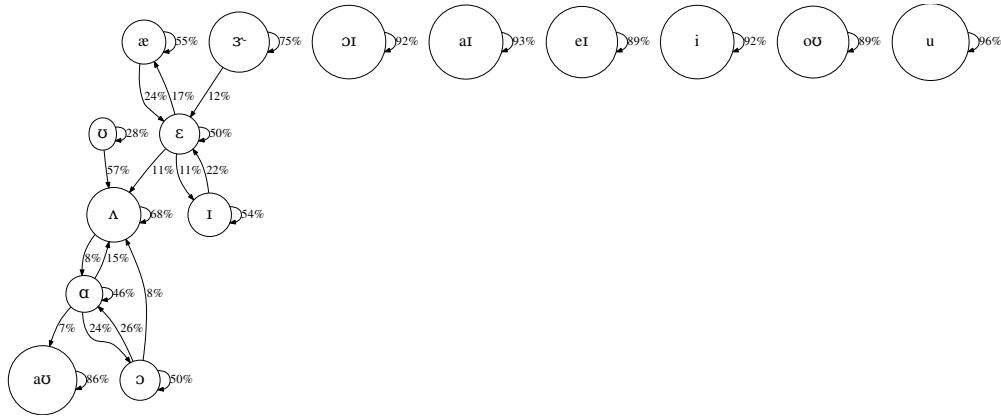
(b) Average RD Phonetic Confusion Patterns for Vowel at Final Position Condition. Confusing centers were /ɛ, ʌ, ɑ/, and /ɔ/. Intact consonants were /aʊ, ɔɪ, aɪ, eɪ, i, oʊ/, and /u/, each with at least 86% correct rate.

Figure 3.23: Average RC and RD Confusion Patterns in VF

61

One confusion pattern ʒ→ʤ (43%) stood out from the others, and its rate of occurrence is also much greater than that in the RC case. Other relatively severe confusion patterns included: θ→ð (22% of the time), θ→f (21%), v→ð (19%), ʤ→tʃ (13%), r→w (12%), ð→v (12%), and v→f (12%). Except for r→w, the above confusion patterns all happened between fricatives and affricates. Less severe confusion patterns were: ʒ→ʃ (9%), ð→θ (9%), v→b (9%), θ→s (8%), z→s (8%), f→v (7%), l→ð (7%), f→θ (5%), f→ð (5%), ð→l (5%), g→k (5%), p→h (5%), and ð→d (4%). Most confusion patterns on this level still happened between fricatives and a few between fricatives and stops.

In summary, the confusion patterns and the groupings among confusing phones were very similar to the RC performance, only with higher rates of confusion. One group was formed by the phones: /ʒ, ʤ, ʃ/ and /tʃ/. All four were palatal-alveolar consonants. Another group consisted of the phones: /ð, θ, f, v, b, s, z, d/ and /l/. Except for /l/, all of these were fricatives, sibilants, and stops. Mostly the confusions were concentrated among fricatives, affricates, sibilants and two stops. Other very different confusion patterns that were not observed in the average RC display were ð→l, l→ð, g→k, r→w and p→h, while z→ʒ in the RC group did not happen here. Since these six patterns had only less than a 10% rate of occurrence, the main structure of all of the RD confusion patterns remained similar to that of the RC's. Centers of confusion were /θ, ð, f/ and /v/.

AVERAGE RD CONFUSION PATTERNS FOR THE CF CONDITION

Figure 3.21b presents RD confusion patterns for the CF condition. Intact consonants were correctly recognized at least 74% of the time. These phones were: /t/, /l/, /d/, /p/, /k/, /tʃ/, and /ʃ/. Difficult phones were: /ð/ (only 42% correct identification), /ʒ/ (55%), /θ/ (64%), /ŋ/ (%), /ʤ/ (68%), /z/ (68%), /f/ (69%).

Confusion patterns that happened at least 10% of the time were: ð→v (25%), ʤ→tʃ (22%), z→s (22%), ʒ→ʃ (18%), ʒ→ʤ(17%), θ→f (17%),f→θ (16%), ŋ→m

(13%), g→k (11%), n→ŋ(11%), ð→θ (10%), and ŋ→n (10%). Other less severe confusion patterns were: v→f (9%), m→n (9%), b→p (9%), r→l (9%), θ→s (8%), s→z (7%), n→m (7%), ʤ→ʒ (6%), v→ð (6%), ʒ→z (5%), f→s (5%), and ŋ→g (4%).

Again, the confusion patterns were similar to those of RC group. When put in final position, nasals and velars /k, g, ŋ, m, n/ formed a confusion category, separated from other phones and confusion patterns. Other very different confusion patterns that were not observed in the average RC confusion patterns were ʒ→z, s→z, b→p and r→l; while patterns like d→t, b→v in the RC group did not happen here. Phones at the center of confusions were /ʤ, ʒ, z, s, f, θ, ð, v, ŋ, n/ and /m/.

AVERAGE RD CONFUSION PATTERNS FOR THE VI CONDITION

Among the four conditions (CI, CF, VI, and VF), the general structure of the average RD confusion patterns for the VI condition differed the most from the structure of the average RC confusion patterns. There were six intact vowels for the RD children: /u/, /eɪ/, /ɔɪ/, /aʊ/, /aɪ/, and /oʊ/, with accuracy rates of at least 71%. However, there was a huge chain of confusion patterns that involved 11 vowels. The difficult phones were: /ɑ/ (only 27% correct identification), /ʊ/ (41%), /ɔ/ (51%), /æ/ (62%), /ɛ/ (67%), /ʌ/ (71%), /ɝ/ (73%), and /ɪ/ (75%).

The most significant confusion patterns were the same as in the average RC display for the VI condition: ɑ→ɔ (47% of the time) and ʊ→ʌ (39%). Other severe confusion patterns were: ɔ→ɑ (23%), æ→ (23%), ɛ→ɪ (13%), ɝ→ɛ (13%), ɔ→ʌ (12%), and ʊ→u (11%), ɑ→ʌ (11%). Confusion patterns that happened less frequently were: i→ɪ (9%), ʌ→æ(8%), and ɪ>eɪ (7%).

In summary, the diphthongs /ɔɪ, aʊ/ and /ɔɪ/ were robust and did not have significant confusion patterns with other phones. No boundary was found to form vowel confusion groups like those for the RC confusion patterns. Phones at the center of confusions were /ʌ, ɛ/ and /ɪ/.

As shown Figure 3.23b, the intact phones for the RD group in the VF condition had accuracy rates of at least 86%. These vowels were: /aʊ/, /ɔɪ/, /aɪ/, /eɪ/, /i/, /oʊ/, and /u/. Difficult phones were: /ʊ/ (only correctly perceived 28% of the time, on average), /ɑ/ (46%), /ɔ/ (50%), /ɛ/ (50%), /ɪ/ (54%), /æ/ (55%), /ʌ/ (68%), and /ɝ/ (75%).

The most severe confusion patterns were: ʊ→ʌ (57% of the time), ɔ→ɑ (26%), æ→ɛ (24%), ɑ→ɔ (24%), ɪ→ɛ (22%), ɛ→æ(17%), ɑ→ʌ (15%), ɝ→ɛ (12%), ɛ→ʌ (11%), and ɛ→ɪ (11%). Less severe confusion patterns were: ʌ→ɑ (8%), ɔ→ʌ (8%) and ɑ→aʊ (7%). In this condition, there was also only one very large group of vowels for all the confusion patterns. Phones at the center of confusions were /ɛ, ʌ, ɑ/ and /ɔ/.

3.9 Clustering by Affinity Propagation

With the data of the raw probability of error and entropy in hand, we can now cluster the children by their performances in these two dimensions.

Fidelity is also known as affinity or euclidean distance (Hopfner, 2014). It is widely used in methods like belief propagation and affinity propagation for data clustering (Frey and Dueck, 2007).

3.9.1 Principle of the Algorithm

The major advantage of using affinity propagation is that users do not need to specify the number of clusters in advance. The algorithm treats each node with equal weight from the beginning of the optimization process, so it does not have bias to begin with for any node in the community. During each iteration, all nodes are visited and each node gathers information from neighboring nodes to calculate

for availability of and responsibility for itself and each neighbor. The node that gets the highest availability would be chosen by the neighbors and reinforced by responsibility. Therefore it emerges as the center of the cluster surrounded by that community. When the information traffic drops to a sufficiently low level, all nodes gravitate towards their own community center. Clusters are then stably formed.

3.9.2 Clustering Results for the NSCM Data

Using the data for probability of error and entropy from each child, clustering for four conditions of the NSCM task were performed [8]. Results are presented in Figure 3.24a, 3.24b, 3.24c, I.1, and 3.24d. Data points with blue text belong to the RD children and data points with black text belong to the RC children.

Children were clustered by their affinity with the two dimensional phonetic perception data, therefore they were not necessarily divided by reading performance. To reveal individual performance on phonetic perception was a major motivation for including the NSCM task in the study. Children in the RD group could have performance similar to the controls' in perceiving phones. In the real learning or intervention process, regardless of the reading level of the individual child, even normal reading children could have difficulty in perceiving certain phones or sounds, just as the children with reading difficulties would be predicted to have. It might be hard to figure out why they have similar confusion patterns, but it is helpful to understand what incorrect perception certain subgroups of children are encountering and to develop effective guidance for accurately perceiving those phones. Customized phonetic education, treatment, or intervention ought to focus on phones instead of reading levels per se.

In Figure 3.24a, the RC children (points with black text) in general had smaller probabilities of error than the RD children, but they were also separated into differ-

---

[8]Open source code was provided online by Frey Lab and Probabilistic and Statistical Inference Group from University of Toronto: www.psi.toronto.edu.

ent groups in terms of entropy, just like the RD children. RC-Evanhad the smallest average probability of error and average entropy for all phones presented. Since no other child performed as well as he did, he formed a cluster all by himself. From his CI directed graph we can see that most of his consonants were correctly recognized. Except for RD-Latisha, all of the other RD children had between one and two confusions for the presented phones on average. Oddly, most of the RC children, except for RC-Evan and RC-Bob, all had more than two confusions for the presented phones on average. Given the two dimensional data in the consonant initial condition, three clusters were formed. The high-error-entropy cluster included RD-Edward, RD-Laura, RD-Shauna, RC-Miguel, RD-Alina, RD-Tony, and RC-Carly. The middle-error-entropy cluster includes RD-Angela, RD-Teddy, RD-Latisha, RD-Norene, RC-Joanna, RC-Bob, and RC-Anton. No one on average perceived consonants in initial position with more than two confusions.

In Figure 3.24b, the clusters again did not separate the RC children and the RD children by their reading abilities. The high-error-entropy cluster only contained RD children: RD-Alina, RD-Norene, RD-Edward, and RD-Latisha. The low-error-entropy cluster only contained RC children: RC-Evan and RC-Anton. However, the middle-error-entropy clusters had crossover children like RC-Carly and RD-Teddy. Most children had two confusions on average. RC-Anton and RD-Laura had approximately one confusion on average for each consonant in final position. RD-Tony, RD-Latisha, RD-Norene, and RD-Alina had more than two confusions for each CF on average.

In Figure 3.24c, except for RC-Anton and RD-Laura, all of the other children had around two confusions for each VI on average. Basically, Anton formed a cluster by his great performance in this VI condition. Laura represented the high-error cluster alone. All of the other children formed the middle-error-entropy cluster.

In Figure 3.24d, except for RD-Alina, all of the other children had less than two confusions for vowel in final position on average. Laura still represented the

high-error cluster alone.

3.10 Individual Ranking Order for Blocking Purpose

In this section, we explore a permutation method to block-diagonalize the confusion matrix. A confusion matrix is a square matrix comprised of perception vectors for all the phones. Each perception vector represents all possible confusion transitions from one phone. If a suitable sequence can arrange the vectors so that they form a block diagonalized matrix, this sequence can help identify the most confusable grouping of certain phones. If a vector A is sequenced to be the neighbor of some other vectors B and C in the diagonalized confusion matrix, then the corresponding phone of vector A is more confusable to the corresponding phones of vector B and C than it is to other phones. The ultimate goal for block-diagonalizing someone's confusion matrix is to identify the grouping of confusable phones for this person.

To block-diagonalize a sparse matrix is similar to narrowing the bandwidth of the matrix. Therefore we convert the problem into narrowing the bandwidth of the confusion matrix. The built-in MATLAB function `symrcm()` is used for narrowing the "bandwidth" of the sparse matrices (Mathworks, 2013), without changing the properties of the sparse matrix (the eigenvalues, the determinant, and the comprising elements).

An example of a 5 by 5 near–symmetric sparse matrix was tested using `symrcm()`. The example here uses an original near–symmetric sparse matrix

$$
\begin{bmatrix}
1 & 3 & 0 & 4 & 8 \\
5 & 7 & 0 & 0 & 2 \\
0 & 0 & 20 & 3 & 0 \\
4 & 0 & 14 & 6 & 0 \\
5 & 0 & 0 & 0 & 3
\end{bmatrix}
$$
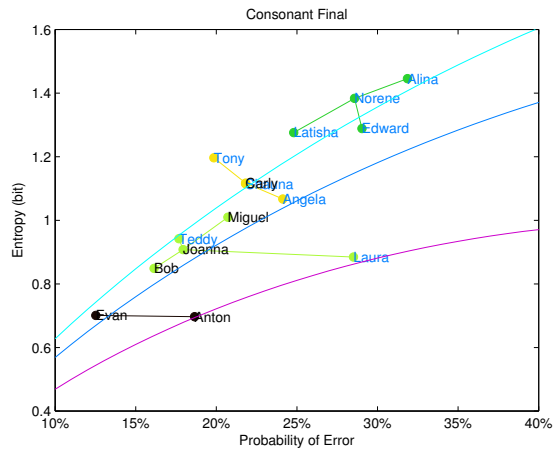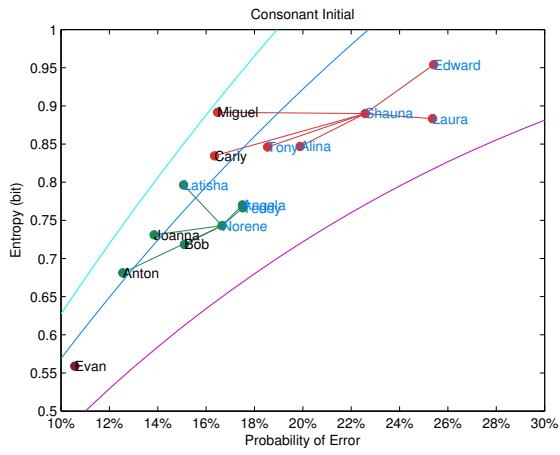
(a) Clusters from Consonant Initial Condition of NSCM task.

(b) Clusters from Consonant Final Condition of NSCM task.
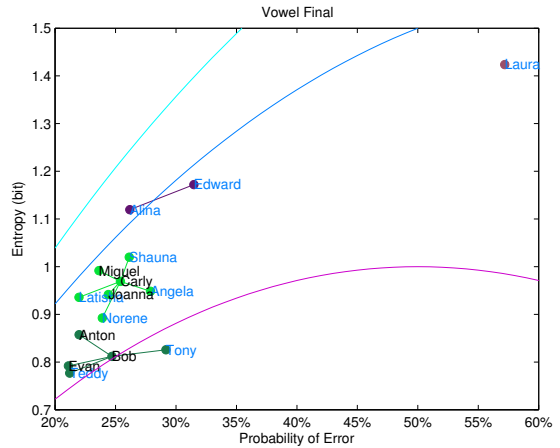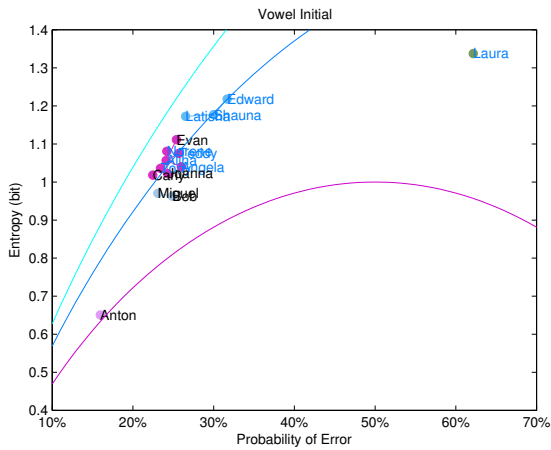
(c) Clusters from Vowel Initial Condition of NSCM task. A zoom-in version for the central cluster please see Figure I.1.

(d) Clusters from Vowel Final Condition of NSCM task.

Figure 3.24: Clustering Results for NSCM. Children with names presented in blue were RDs. Children with names presented in black were RCs. Clusters are connected inside themselves by links. The top cyan curve in each panel is the three-confusion-per-phone boundary. The middle blue curve is the two-confusion-per-phone boundary. The bottom blue curve is the one-confusion-per-phone boundary. Therefore, in most cases, on average, children have one to three confusions for each phones. Many RD children performed similarly as RC children, when they are assessed from both dimensions (Probability of error and number of confusions per phone).
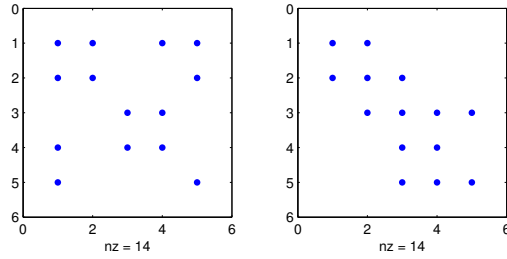
Figure 3.25: `symrcm()` example: the left figure represents the original sparse matrix, the right figure represents the matrix after permutation. As we can see, after the permutation, the nonzero elements get closer to the diagonal. Therefore, vectors which have most information traffic between become neighbors. The permuted sequence of the vectors helps us identify the most information exchanged vectors for each vector in the matrix.

After the permutation (3,4,1,5,2) of rows and columns to minimize the bandwidth, the resulting matrix becomes

$$
\begin{bmatrix}
20 & 3 & 0 & 0 & 0 \\
14 & 6 & 4 & 0 & 0 \\
0 & 4 & 1 & 8 & 3 \\
0 & 0 & 5 & 3 & 0 \\
0 & 0 & 5 & 2 & 7
\end{bmatrix}
$$

The visualization of the two sparse matrix is shown in Figure 3.25. As we can see, blocks are formed around the diagonal elements following permutation.

For a more thorough demonstration of this method, a series of random sparse matrices having different density (ratio of nonzero elements to all elements) levels were generated (Figure 3.26). When the density of the elements in the matrix increases, the diagonalization effect decreases. Even though the confusion matrices are not mathematically symmetric ($A = A^T$), they are near–symmetric. The algorithm treats the input as the sum of the original confusion matrix and its transpose to get a symmetric matrix then minimize the bandwidth of this matrix to get the

69

(a) 10% density

(b) 20% density

(c) 30% density

(d) 40% density

(e) 50% density

(f) 60% density

(g) 70% density

(h) 80% density
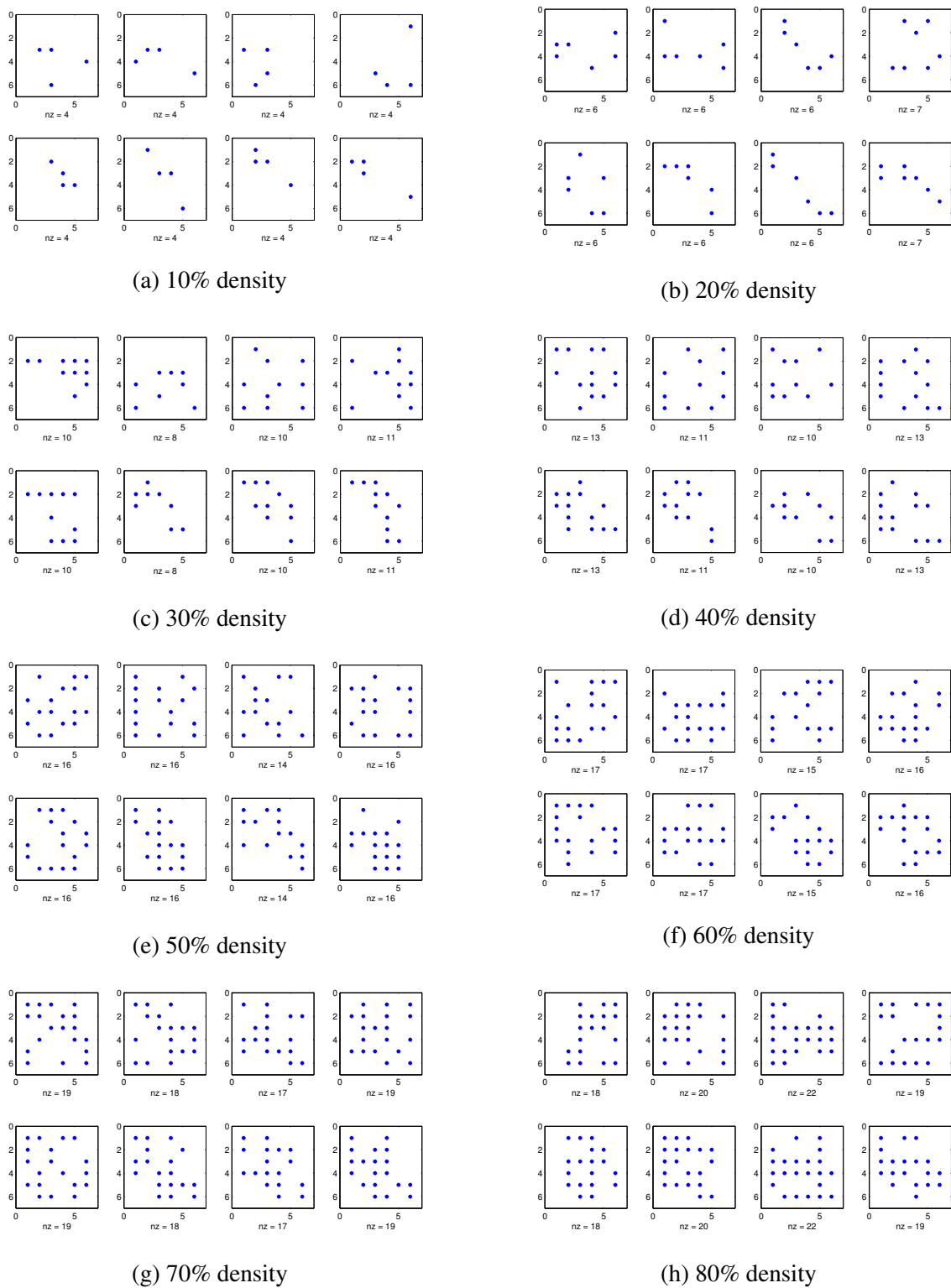
Figure 3.26: In each subfigure, the upper row has the original sparse matrices, and the lower row has the permuted matrices. Note that matrices with smaller density have greater decrease in bandwidth.

permutation result. The results for RD and RC children are shown in Figure F.1 and Figure F.2.

Compared with the directed graphs, the automatic permutation sequencing to identify confusion groups of phones for each subject has unstable performance over different confusion matrices. Still, some good cases by permutation sequencing are Anton CI, Edward CI, Latisha CI, Tony CI, Tony CF, Miguel VI, Norene VI, Shauna VI, Teddy VI, Tony VI, Anton VF, Evan VF, Latisha VF, and Shauna VF.

3.11 Eigenvalues to Reorder Stochastic Transition Matrix

Another way to block-diagonalize the confusion matrix reorders the ranking of phones and is explored with the eigenvalues of the confusion matrix.[9]

For example, a confusion matrix for presented phones [/p/,/t/,/k/,/f/]

$$
C = \begin{bmatrix} 1/2 & 0 & 0 & 1/2 \\ 1/4 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 1/4 & 0 & 0 & 1/2 \end{bmatrix} = \begin{bmatrix} c_1 & c_2 & c_3 & c_4 \end{bmatrix}, \tag{3.4}
$$

where c_k is the corresponding *probability vector* for each presented phone. Because each column is a probability vector, it must sum to one.

The first row of $c_1(1) = 1/2$ specifies the probability of $/p/$ being identified correctly, which is $1/2$. The second row of $c_1(2) = 1/4$ specifies the probability of $/p/$ being reported as $/t/$. The second row of $c_2(2) = 1$ says that when /t/ is heard, it is recognized correctly 100% of the time. Likewise for /k/, described by $c_3(3) = 1$).

Figure 3.27 is a *state diagram* for confusion matrix C. It is also called a *directed graph*. The first state is represented by the circle labeled /p/. It shows that when

---

[9]This section is completely copied from personal communication materials with Professor Jont Allen.
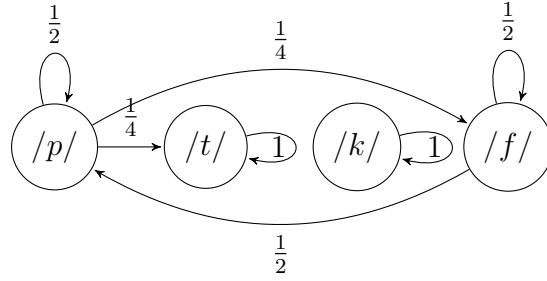
Figure 3.27: *A directed graph corresponding to the Markov matrix of Eq. 3.4.*

/p/ is presented, /p/ is reported 1/2 the time, /f/ is reported 1/4 of the time, and the remaining 1/4 of the time, /t/ is reported. When either /t/ or /k/ is heard, it is always reported correctly (loop-back probability = 1). When /f/ is heard, 1/2 the time it is reported correctly, and 1/2 the time /p/ is reported. If one enters this graph with /p/ presented, eventually one ends up with /t/ as the most probable response state. This is due to the fact that once /t/ is entered, there is no escape. Given enough trials, /t/ is the most probable state, since once it is entered, there is no exit condition.

The ending states depend on the initial state probability conditions π_0. If we present /p/ n times ($\pi_0 = [1, 0, 0, 0]^T$), we may determine the expected outcome as $\pi_n = C^n \pi_0$, with $\pi_0 = [1, 0, 0, 0]^T$.

3.11.1   A stochastic experiment

For example, the *maximum entropy* starting condition is when the apriori probabilities are all equal. In that case $\pi_0 = [1, \ 1, \ 1, \ 1]^T/4$. Alternatively, the lowest entropy starting condition is, for example, $\pi_0 = [1, \ 0, \ 0, \ 0]^T$, corresponding to /p/ being presented on the first trial. Each of these two cases is interesting to explore.

The first trial in our probability experiment starts from the maximum entropy state, which renders the probabilities

$$\pi_1 = C\pi_0 = \frac{1}{4}\sum_{p=1}^{4} c_p,$$

72

i.e., $\pi_1 = 1 = \sum[0.25, 0.3125, 0.25, 0.1875]^T$. What this tells us is that if you present sounds to a person with a confusion matrix given by Eq. 3.4, they will have a bias for reporting [/p/, /t/, /k/, /f/] of π_1. The probability of /p, k/ stayed the same, while the probability of /t/ went up by 1/16, and the probability of /f/ went down by 1/16. Namely the row sums did not change, while the column sums are all 1. After two trials, the bias becomes

$$\pi_2 = C\pi_1 = C^2\pi_0 = [0.2188, 0.375, 0.25, 0.1562]^T$$

After a third iteration, we find

$$\pi_3 = C^3\pi_0 = [0.1875, 0.4297, 0.25, 0.1328]^T.$$

$\pi_{10} = C^{10}\pi_0 = [0.0726, 0.6261, 0.25, 0.0513]^T$. After 100 iterations, the row sums have converged to $[0, 0.75, 0.25, 0]$. In summary

$$C^{100} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \quad \text{and} \quad C^{100}\pi_0 = \begin{bmatrix} 0 \\ 0.75 \\ 0.25 \\ 0 \end{bmatrix}$$

The key is to study C^n, as an eigen matrix decomposition. The eigenvectors and eigenvalues for C^n are

$$E = \begin{bmatrix} 0 & 0.476 & -0.794 & 0 \\ 1 & -0.812 & 0.233 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0.337 & 0.561 & 0 \end{bmatrix}, \quad \lambda = \begin{bmatrix} 1 \\ 0.854 \\ 0.146 \\ 1 \end{bmatrix}$$

From this analysis we see that eigenvectors 1 and 4 have eigen values of 1, which says they define stationary states of our CM. Namely if state 1 (/p/) is spoken, the final state is 2 (/t/). Likewise if state 4 (/f/) is spoken, the final state is 3. The final conclusion for this experiment is that, by sorting the eigenvalues, we can

get the ranking for the four presented phones by their transition traffic.

Using this method, some good examples in sequencing the phones from confusion matrices are: Anton VI, Bob VF, Carly VF, Edward VF, and Teddy VF.

Chapter 4

DISCUSSION

Research questions proposed in the introduction will be addressed in this section based on results presented in the last chapter.

4.1 Did RD Children Perform Worse Than Controls?

Some studies claimed small or no significant speech perception deficit was found in dyslexia children (Hazan and Adlard, 1998) (Ziegler *et al.*, 2009) (Messaoud-Galusi *et al.*, 2011). However, from our results, we conclude that notable phonetic speech-perception deficit exists for the children with reading disabilities compared with children with normal reading ability. On average, probability of error was 6.56% for RC and 20.77% for RD.

As shown by the sorted error plots and repeated measure testing results from Section 3.2 and 3.5, in general, the RC group performed significantly better than the RD group on both the SCO and NSCM tasks. The RC group had more zero error phones and lower probabilities of error for nonzero error phones than the RD children on the SCO task. Though statistical testing affirmed that significant differences exist between the two groups for both the SCO and NSCM tasks, only the SCO results were significantly correlated with a number of measures of reading ability, for all of the children combined.

The results of the clustering of children on the NSCM tasks (Figure 3.24) did not reveal a clear separation of the two groups. Nevertheless, the RC children generally tended to be located in the lower region of entropy for all four conditions

(CI, CF VI, VF) than the RD children. In contrast, some of the biggest clusters that occupied the middle range entropy consisted of children from both groups.

4.2 Main Effects and Significant Interaction for Both Tasks

The mean values of each main effect and significant interaction in the linear models are listed in appendix D.

SCO results interpretation: The statistic test showed significant difference in subject group (RD / RC) and phone type (consonant / vowel). The RD group had higher error than the RC group, and the consonants had higher error than the vowels.

Interaction between subject group and phone type also showed significant effect. For the consonants, the RD group had higher error than the RC group. For the vowels, the RC group had higher error than the RD group.

Interaction among subject group, phone position and phone type showed significant effect too. In CF condition, the RC group had higher error than the RD group. In CI condition, the RD group had higher error than the RC group. In VF condition, the RD group had higher error than the RC group. In VI condition, the RC group had higher error than the RD group.

Detailed data please refer to appendix D.

NSCM results interpretation: The statistic test showed significant difference in subject group (RD / RC), phone position (syllable initial / syllable final) and phone type (consonant / vowel). The RD group had higher error than the RC group; the syllable final position phones had more error than the phones in the syllable initial position, and the vowels had higher error than the consonants.

Interaction between subject group and phone type also showed significant effect. For the consonants, the phones in syllable final position erred more than

the phones in syllable initial position. For vowels, the phones in syllable initial position erred more than the phones in syllable final position.

4.3   Comparison of Tasks

Besides a conclusion that RCs performed generally better than RDs in these non-categorical phonetic perception tasks, the results from the previous chapter promote a comparison of tasks to offer more insight into the phonetic perception related to reading disabilities.

4.3.1   Task Effectiveness

By and large, the SCO task yielded much larger differences between the RD and RC groups than the NSCM task. This result is consistent with Hazan and her colleague's findings that performance differences were more consistent for discrimination than for identification (Hazan *et al.*, 2013). Similar results were observed in another study by Hazan and Adlard (1998).

The phonetic perception process was engaged, perhaps even taxed, during the SCO task. The children were not told whether the change of phone was in the syllable-initial position or in the syllable-final position. They were presented with utterances with the only instructions being that they would hear three syllables and then should find the one that had a change in sound structure. The SCO task is presumably a more difficult task than the NSCM task, since it involves comparisons and permutations of sounds in syllable triads. Nonetheless, normal reading children had less error on this task compared to their results on the NSCM.

One explanation for the task difference might be related to pre-cognitive processing, that is, auditory memory is required when children need to do permutation comparisons of the three phones in order to identify the two equal phones and then point out the oddball. As seen from the results of both tasks in Figures 3.2

and 3.8, there exists a notable difference in the performance of the two groups. It seems likely that the children with RD struggled with mismatched phone representations during the SCO task, requiring greater effort to store all the information they thought they heard in the temporary auditory memory system, instead of processing the phones easily and correctly right away. Some of the RC children had the same error magnitude, but in general as a group, they outperformed the RD children. This difference was not as obvious in the NSCM task. Hence, more experiments on the SCO task with various numbers of potentially confusing phones (e.g. a triad, quartet, or quintet of syllables) would be needed to determine if working memory specific for auditory processing is the source behind its contrast from the NSCM task. A recent meta-analysis (Fullgrabe and Rosen, 2016) concluded that less than 2% of the variation in speech perception is contributed by working memory for normal hearing young children. Here the error difference between tasks for the RC children is also less than that for the RD children. This could suggest that greater effort was needed for the RD children to hold the three phones together on the SCO task. The RC children appeared to have fewer issues with auditory processing and memory.

4.3.2 Task Informativeness

From the experiment results, the SCO task was more effective in separating the two groups, and had the possibility of revealing phonetic memory as an important variable in reading ability. However, to improve the RD intervention, pinpointing low performance is key. While SCO is more sensitive, NSCM has proven to be more informative in providing the necessary information for intervention.

Bench-marking children with normal reading levels could be best obtained for reference using NSCM. Any deviation from normal patterns of confusion that are observed in children with RD could be viewed as barriers to normal phonetic perception. Hence training based on this extra directed graph information may be

offered to overcome difficulties in decoding these confusing phones.

Inside the NSCM task, performance in general was better in the CI and VF conditions than in CF and VI conditions. A possible explanation could be that the most frequent syllables in English are CVs, thus our phonetic perception in these situations are likely stronger than for VC and VCV sounds (Allen, 2005).

4.4 Similarities and Differences Between the Two Groups for Phonetic Perception

After the review of the design advantages of each task, we can now summarize some facts about phonetic perception for each group of subjects.

4.4.1 Similarities between the Two Groups of Children

The analysis of average confusion of individual phones for both RC and RD groups from Section 3.8 showed us their similarities in phonetic perception:

Consonants   1. Both the RD and RC children had great difficulty in correctly perceiving consonants like some affricates and fricatives in syllable-initial position. The most frequent confusions were ʒ→ʤ, θ→ð, θ→f and v→ð. Notable phones at the center of confusions were /θ/, /ð/, /f/, and /v/.

Two major categories of confusing phones were formed. The first category includes palatal-alveolar obstruents: /ʒ/, /ʃ/, /tʃ/, and /ʤ/. The second category includes anterior obstruents /z/, /s/, /θ/, /ð/, /v/, /f/, and /b/.

2. For consonants in syllable-final position, frequent confusion patterns for both the RD and RC children were f→θ, ŋ→n, ð→v, θ→f, z→s. Common high entropy phones for both the RD and RC children were /θ/, /ð/, /s/, /v/, /m/, /n/ and /ŋ/. Besides the two major categories that exist in the CI condition, both the RD and RC children had another confusion category formed by nasals and velars:

/ŋ/ , /m/, /n/, /g/ and /k/. From the RC benchmark we know that in English some consonants articulated by talkers in syllable-final position may be poorly articulated, so that listeners' perceptions may be weak for consonants in final position. Entropies calculated from the CF condition were greater than those from other cases, showing a tendency toward guessing by the RD children (see Figure 3.14d).

Vowels 3. For both the RD and RC children, there were also two major categories of confusing vowels. The first category includes the front vowels: /æ/, /ɛ/, /ɪ/, and /i/ (with /i/ only in the VI condition, the other three vowels in both the VI and VF conditions). The second category includes the back vowels: /ʊ/, /ɑ/, /ɔ/, and /ʌ/. Coincident or not, phones in the same category also frequently correspond with their common allophones for certain spelling patterns.

4. Diphthongs like /ɔɪ/, /aʊ/ and /aɪ/ and tense vowels like /oʊ/, /eɪ/ and /u/ are not confusable for either the RD or the RC children, regardless of position (initial, final). The tense vowel /i/ could be tacitly recognized by both the RD and RC children only when it was presented in the syllable-final position.

5. Severe vowel confusion patterns that are common to both the RD and RC children were: ʊ→ʌ, ɑ→ɔ, ɔ→ɑ, and æ→ɛ. Notice that the second and the third confusions are common in Midwestern dialect in American English.

4.4.2 Dissimilarities of the RD group

From comparison of the data between the two groups, observable distinctions of the RD group can be made as follows:

Consonants 1. In the CI condition, specific patterns, such as confusion between liquids and fricatives (l→ð and ð→l) or stops and fricatives (p→h) and voicing confusions (g→k) uniquely belonged to the RD group (see Figure 3.20b).

2. In the CF condition, the RD children possessed special patterns such as ʒ→z, s→z, b→p, and r→l (see Figure 3.21b).

Vowels   1. In the VI condition, the RD group had a confusion pattern that did not exist for the average RC group : ʌ→æ, and this pattern linked the two major categories that existed in the average RC group into one big cloud of confusions. Also, the RD group had unique branched-out patterns where they tended to recognize lax vowels as tense vowels or vice versa, such as: ʊ→u, ɝ→ɛ, and ɪ→eɪ (see Figure 3.22b).

2. In the VF condition, unique patterns belonging only to the RD children were ɝ→ɛ and ɑ→aʊ. Additionally, while the transition link between the two major categories for the RC children was æ→ɑ, the link between the two categories (front or back vowels) for the RD children was ɛ→ʌ (see Figure 3.23b).

4.4.3   Potential Reasons

After the detailed consideration of similarities and distinctiveness of the confusion patterns, it is necessary to look at the potential reasons that contribute to these perceptual patterns, as they relate to reading ability.

One perspective for answering why the confusing categories formed the way they did in these results is to try to look at the acoustic or perceptual cues for these confusing phones. Li reported the perceptual cues for 12 consonants in his PhD thesis (Li, 2009). He provided a graphic summary of perceptual cues in Figure 4.1. The phones in our first consonant confusion category, /ʃ, ʒ, tʃ/, and /dʒ/ locate at an extended area in the spectrogram, 2-8 kHz in the frequency region and from around 160 ms before the following vowel to the start of vowel. Our second confusion category for consonants has /f, v, θ, ð, s/ and /z/. The region on the spectrogram for /f/ spans from 0.8 to 2.8 kHz in frequency and about 80 ms in duration before vowel. The region on the spectrogram for /v/ spans from 0.5 to
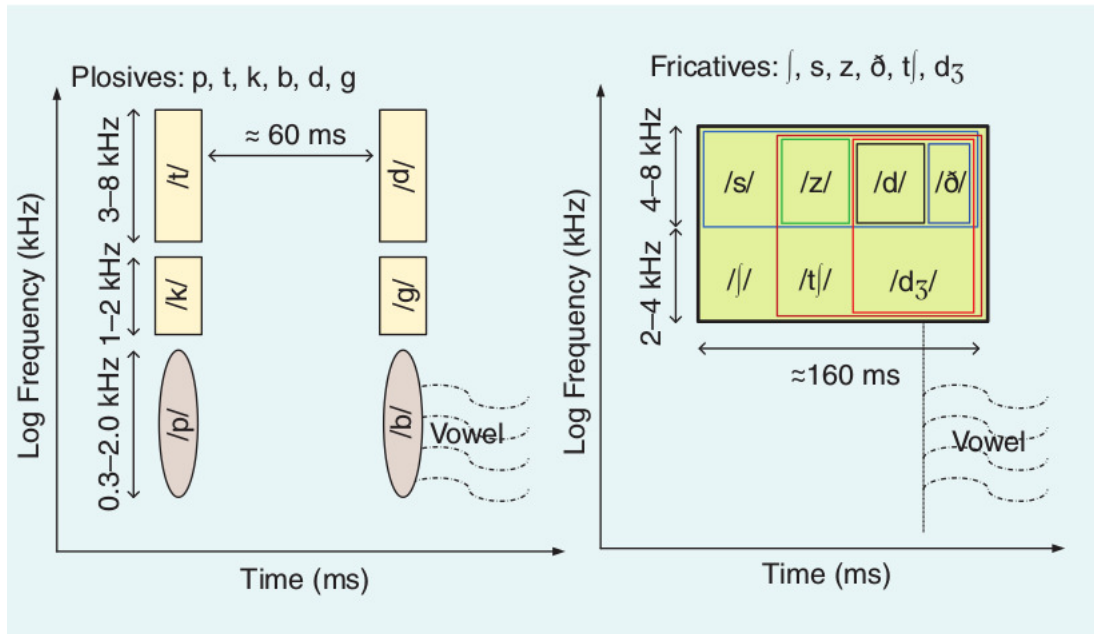
Figure 4.1: A schematic drawing of the perceptual cues for initial consonants preceding vowel /a/, in terms of time-frequency allocation (Li, 2009).

1.4 kHz in frequency and about 40 ms in duration. The phones /θ and ð / do not have a strong and compact perceptual cue, hence can be confused with many other phones. The region on the spectrogram for /s/ spans from 4 to 8 kHz in frequency and about 100 ms in duration before vowel. The region on the spectrogram for /z/ spans from 3 to 7.5 kHz in frequency and about 50-70 ms in duration. Therefore for the six consonants in this second category, /f/, and /v/ are close to each other; /s/ and /z/ are close to each other; /θ/ and /ð/ function as the linchpin which builds the confusion links among them. Our third consonant confusion category contains /g, k, m, n/ and /ŋ/. The phone /g/ is characterized by an energy burst at the frequency region of 1.4-2 kHz lasting about 40 ms in duration. The phone /k/ is characterized by an energy burst at 1.6 kHz lasting for about 50-70 ms in duration. The two are close in these perceptual cues and confusable. On the other hand, /m/ and /n/ have F2 formant transitions at around 1kHz, both of which last for about 100ms.

The major two categories of confusion in vowels can possibly be explained with the vowel quadrilateral (Reetz and Jongman, 2008). The first category contains

the front vowels /æ/, /ɛ/, /ɪ/, and /i/, which mainly have the first formant frequency ranging from 0.3 kHz to 0.7 kHz and the second formant frequency ranging from 1.6 kHz to 2.6 kHz. The second category contains the back vowels /ʊ/, /ɑ/, /ɔ/, and /ʌ/, which mainly have the first formant frequency ranging from 0.45 kHz to 0.9 kHz and the second formant frequency ranging from 0.9 kHz to 1.5 kHz. These phones are confusing for normal listeners and readers, not to mention children with reading disabilities who might have phonetic perceptual deficits.

The second perspective is that the perception of the four consonants /ʒ/, /ʃ/, /tʃ/, and /dʒ/ in this first confusing category may mature later than that of other consonants in English. Words containing these phones tend to be acquired later in children's speech development. When learning to read, normally developing children start by pronouncing phones of the word. Frequently appearing CI phones like /p/, /g/, /l/ could significantly contribute to the decoding process during reading. Correctly perceiving them provides an important aid in the reading process. The 'Late Eight' consonants /θ, ð, s, z, ʃ, ʒ, r/, and /l/ are considered to be correctly acquired later than other consonants for children (Shriberg and Kent, 1982).

A third perspective or potential reason for RDs' dissimilarities from RCs could be the duration or pitch (prosodic characteristics) of the syllable varing from the prototypical sound. RC children may easily map the presenting syllable to certain parts of their inner library of known pronunciation patterns. The mature connections between the acoustic sounds and their perceptual (abstract) representation will obviously be important to the performance. For the RD children, unfamiliar phones or stress of the phones that did not match any remembered sound patterns would likely lead to perceptual errors.

In summary, the reasons for phone perceptual similarities and distinctions could be differences in the salience of perceptual cues, less maturity in certain late acquiring phones, talker characteristics of the stimuli, and other factors.

4.5 Outliers in Our Phonetic Perception Tasks

Though on both the SCO and NSCM tasks the RC group performed significantly better than the RD group according to the analysis, not all RDs performed worse than RCs. By close inspection of Figure 3.2, on the SCO task, Norene and Teddy were performing at the average RC level for Consonants in Initial position, Shauna and Teddy were at the average RC level for Consonants in Final position and Vowels in Initial position, and Shauna again was at the average RC level for Vowels in Final position. Teddy was also at the average RC level for the four conditions in the NSCM task (see Figure 3.8). Thus, these outliers in the RD group all demonstrated normal phonetic perception abilities similar to the normal reading children. In the Post Hoc analysis (Table 3.2 and Table 3.5) where these RDs were removed, the performance separation between groups became more significant. A recent study (Corriveau *et al.*, 2007) also pointed out that, in their experiment, only 70-80% of the children with specific language impairment (a disorder similar to and sometimes co-morbid with dyslexia or reading disability) showed deficits in auditory processing skills.

In conclusion, the existence of these RDs with good phonetic perception indicated factors other than weakness in phonetic perception could be contributing to their reading problems .

Chapter 5

FUTURE WORK

5.1   Fidelity, Hellinger Distance and Log Scale

By definition, a probability vector $\vec{p_i}$ is a vector of probabilities for multinomial distributed outcome events with n = 24 or 15 from a single source event i. Taking the square roots of the original probability vectors $\sqrt{\vec{p_i}}$ can help us construct a complete and ortho-normal vector space while maintaining the distribution property of the data so as to establish the problem in a confined domain. One good thing we can get from a complete and ortho-normal vector space is that now the probability vectors can be compared on the same scale with the same ruler and it is a robust way to detect furthest distance deviation, which is the emergence of abnormal signal. Another benefit is that the square root transformation could amplify the smaller probability values region. With that being said, fidelity is defined as the inner product of two square rooted probability vectors. A special example is that the summation of probabilities of all states equals one:

$$(\sqrt{\vec{p_i}}, \sqrt{\vec{p_i}}) = \sum_i p_i = 1.$$

Since they are already defined with norm one, the fidelity will satisfy the Cauchy-Schwarz inequality and can be viewed with the physical meaning as the cosine of the angle between two square rooted probability vectors. Hellinger distance is then defined as one minus the fidelity, yielding value one to identify the existence of closest possible similarity between two vectors and zero as two vectors being perpendicular with each other, which literally means no similarity between the two

in this defined vector space. Other nice features of a complete normed space are not used here yet, but they could be of value in future analysis.

The fact that most of the data are being presented in log scale also comes from the need to emphasize the smaller probability region and stabilize the variance.

5.2 The Correlation between Phonetic Perception in SCO and Phonetic Perception in NSCM

It would be interesting to examine the perception of every phone between SCO and NSCM. Since the two tasks differ in the extracted aspects of phonetic perception, if there is a close correlation between the scores for specific phones of the two tasks that could define the separation of RD from RC, the two tasks could be merged into one to identify children with RD and their confusion patterns with much less testing time.

We had already extracted the average patterns from children with normal reading ability. For phones in each situation (CI, CF, VI, and VF), there were boundaries between very robust phones and easily confusable phones for RCs (see Figure 3.20a, 3.21a , 3.22a, 3.23a).

5.3 Time Variables

Variation in different time variables might play a role in perceiving phones.

The first time variable is the duration time of the target phone. For example, in NSCM, RD had higher error and entropy for consonants in final position than in initial position. A major difference in the perceptual duration of the consonants was observed. It is helpful to understand if probabilities of error correlate with the duration of the target phone. For CI, this time starts from the beginning of onset and ends at the beginning of rime. For CF, this time starts from the end of rime

and ends at the end of the syllable.

The second time variable is the response time for target phone. A phone could be correctly or wrongly perceived, but if it took too long for the child to perceive, it is surely a speech perception problem. The presented phone might not have formed a firm and correct phonetic representation inside the brain. Whether it could become a source for reading disabilities for some children needs investigation.

5.4 Talker Effects

A talker effect was reported in some studies as making a contribution to error patterns. To investigate if this variable had an effect on our experiments, phones should be categorized with a tag on their producers (i.e., an indication of who was the talker). Correlations between the individual talker and task receiver should be characterized. If conditions permit, each phone category from an individual talker should also be characterized by his or her subsequent error patterns to see if the current findings confirm those of other studies. If this factor plays an important role in perceptual error patterns, it should be carefully controlled in future experimental designs.

5.5 Standard Deviation Trend

A summary for SCO task performance for all subjects, both kinds of phones and at both positions is presented in Table 5.1. Standard deviation of RC subjects error rates converged after 5 to 6 trials, as plotted in Figure 5.1. Therefore in the future, more than 5 to 6 trials should be carried out to test for phone perception for RC subjects.

Table 5.1: Convergence table for all variables.

trial#	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
p_rc	0.073	0.074	0.068	0.066	0.065	0.065	0.065	0.065	0.065	0.066	0.066	0.067	0.067	0.067	0.067	0.068	0.068	0.069	0.069	0.068
sd_rc	0.101	0.101	0.081	0.067	0.063	0.062	0.062	0.062	0.060	0.060	0.060	0.060	0.060	0.060	0.060	0.059	0.058	0.058	0.057	0.057
p_rd	0.203	0.205	0.204	0.204	0.207	0.209	0.209	0.210	0.211	0.208	0.207	0.210	0.204	0.201	0.201	0.205	0.203	0.199	0.198	0.192
sd_rd	0.160	0.160	0.158	0.157	0.156	0.154	0.147	0.147	0.146	0.145	0.144	0.145	0.140	0.138	0.138	0.133	0.132	0.132	0.131	0.123
p_f	0.170	0.174	0.170	0.167	0.167	0.166	0.162	0.162	0.159	0.154	0.153	0.152	0.146	0.145	0.145	0.140	0.140	0.139	0.134	0.127
p_i	0.135	0.135	0.135	0.134	0.135	0.136	0.135	0.135	0.135	0.135	0.132	0.133	0.127	0.122	0.122	0.123	0.119	0.116	0.114	0.114
p_c	0.154	0.156	0.153	0.150	0.150	0.150	0.146	0.145	0.144	0.137	0.132	0.130	0.124	0.117	0.117	0.113	0.107	0.105	0.103	0.099
p_v	0.150	0.150	0.150	0.150	0.149	0.150	0.151	0.152	0.151	0.153	0.154	0.156	0.152	0.152	0.152	0.153	0.154	0.152	0.148	0.143
sd_f	0.169	0.169	0.161	0.156	0.155	0.153	0.148	0.149	0.148	0.145	0.143	0.144	0.137	0.136	0.136	0.130	0.128	0.127	0.125	0.117
sd_l	0.136	0.136	0.135	0.134	0.133	0.132	0.127	0.127	0.125	0.124	0.121	0.122	0.118	0.112	0.112	0.109	0.105	0.104	0.100	
sd_c	0.161	0.161	0.154	0.149	0.148	0.144	0.135	0.135	0.133	0.127	0.123	0.122	0.117	0.111	0.111	0.110	0.103	0.100	0.099	0.093
sd_v	0.143	0.143	0.142	0.142	0.142	0.142	0.142	0.143	0.142	0.143	0.143	0.144	0.137	0.137	0.137	0.131	0.131	0.129	0.127	0.121
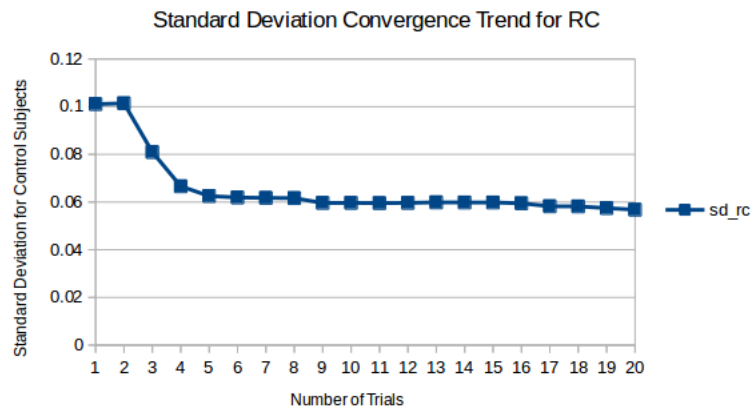
Figure 5.1: Standard deviation convergence plot for RC.

Appendix A

TABLE A.1 FROM DARPABET TO INTERNATIONAL PHONETIC ALPHABET

Table A.1: LDC unvoiced consonants, voiced consonants and vowels.

Unvoiced Consonants			Voiced Consonants			Vowels			
Example	Dbet	IPA	Example	Dbet	IPA	Example	Dbet	IPA	L/T
/ch/urch	C	tʃ	/th/is	D	ð	h/a/t	@	æ	
/sh/e	S	ʃ	/b/ee	b	b	h/u/t	A	ʌ	T
/th/ink	T	θ	/d/og	d	d	b/e/t	E	ɛ	L
/f/ish	f	f	/g/ab	g	g	h/i/t	I	ɪ	L
/h/e	h	h	/j/udge	J	ʤ	b/oy/	O	ɔɪ	
/c/at	k	k	/l/ook	l	l	b/ir/d	R	ɝ	
/p/en	p	p	/m/an	m	m	p/u/t	U	ʊ	L
/s/ee	s	s	/n/ap	n	n	h/ow/	W	aʊ	
ca/t/	t	t	/r/eal	r	r	wh/y/	Y	aɪ	
/wh/at	H	ʍ	plea/s/ure	Z	ʒ	c/au/ght	c	ɔ	T
			si/ng/	G	ŋ	f/a/ther	a	ɑ	
			/v/ow	v	v	b/ai/t	e	e / eɪ	T
			/w/in	w	w	b/ee/	i	i	T
			/y/ou	y	j	b/oa/t	o	o / oʊ	T
			/z/oo	z	z	b/oo/	u	u	T
						th/e/		ə	L

Taken from http://jontalle.web.engr.illinois.edu/Public/Corpus/LDC_symbols.pdf

Appendix B

TABLE B.1 READING SCORES FOR RD AND RC GROUPS

Table B.1: Reading, IQ, and Production Scores for RD and RC groups.[1]

	WI	WA	R-FLU	R-COMP	KBIT-M	PPVT-III	GFTA-2
Reading Control Group							
Mean	112.33	112.83	14.50	14.33	119.5	119.67	104
SD	10.46	7.96	4.32	2.58	8.89	15.85	1.67
Reading Disability Group							
Mean	90.00	90.89	6.67	8.22	102	94.44	98.67
SD	8.59	8.18	2.45	3.31	13.91	16.57	8.99
Welch T-Test Results[2]							
t value	4.34	5.18	4.03	4.64	2.9715	2.9653	1.736
p-value	0.0017	0.00029	0.0048	0.00049	0.01	0.01	0.12
df	9	11	7	13	13	11	9

[1] The Word Identification (WI), Word Attack (WA) measures came from the Woodcock Reading Mastery Tests-Revised and were measured according to standard scores (Mean = 100, SD = 15). The Reading Fluency (R-FLU) and Reading Comprehension (R-COMP) measures came from the Gray Oral Reading Test 4th Edition, and were measured according to scaled scores (Mean = 10, SD = 3). The Kaufman Brief Intelligence Test: Matrices subtest (KBIT-M) was used to measure the non verbal intelligence of the children. The Peabody Picture Vocabulary Test 3rd Edition (PPVT-III) was used to measure the verbal intelligence of the children. The Goldman-Fristoe Test of Articulation 2nd Edition (GFTA-2) was used to measure the articulation ability of the children.

[2] Welch t test was used for testing because variances of the measures for the RC and RD children were different. The degrees of freedom was calculated as $df = \frac{(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2})^2}{\frac{(\frac{s_1^2}{n_1})^2}{n_1 - 1} + \frac{(\frac{s_2^2}{n_2})^2}{n_2 - 1}}$ (Pace, 2012).

Appendix C

SUBJECT INFORMATION

Table C.1: Originally there were 19 subjects. Data of four subjects were left out because of lack of data: Savannah and Lucas barely started the study and only did a few sessions, Matt and Tina only completed the SCO task but not the NSCM.

Pseudonym	Acronym	Group	Age
Anton	At	RC	11;4
Bob	Bb	RC	9;10
Carly	Cl	RC	8;9
Evan	Ev	RC	11;6
Joanna	Jn	RC	10;3
Miguel	Mg	RC	10;6
Alina	Al	RD	10;8
Angela	Ag	RD	9;0
Edward	Ed	RD	8;5
Latisha	Lt	RD	8;5
Laura	Lr	RD	9;11
Norene	Nr	RD	9;10
Shauna	Sn	RD	10;1
Teddy	Td	RD	8;4
Tony	Tn	RD	9;0

Appendix D

STATISTICS OF MAIN EFFECTS AND SIGNIFICANT INTERACTIONS FOR BOTH TASKS

Table D.1: Main Effects for SCO

	Subject	Phone Position	Phone Type
Category	RC	Initial	Vowel
Mean	0.0656	0.1346	0.1528
StandardDeviation	0.5966	0.1355	0.1426
Category	RD	Final	Consonant
Mean	0.2077	0.1542	0.1536
StandardDeviation	0.1601	0.1686	0.1606
Conclusion	RD $>$ RC	Final$>$Initial	Consonant $<>$ Vowel

Table D.2: Main Effects for NSCM

	Subject	Phone Position	Phone Type
Category	RC	Initial	Vowel
Mean	0.1877	0.2132	0.2721
StandardDeviation	0.1743	0.2065	0.2447
Category	RD	Final	Consonant
Mean	0.2535	0.2412	0.1992
StandardDeviation	0.2205	0.2051	0.1715
Conclusion	RD $>$ RC	Final$>$Initial	Vowel $>$ Consonant

Table D.3: Significant Interaction for NSCM

	PhonePosition:PhoneType	
	type=con	type=vow
Category	Final	Final
Mean	0.2227	0.2709
StandardDeviation	0.1636	0.2556
Category	Initial	Initial
Mean	0.1756	0.2732
StandardDeviation	0.1761	0.2338
Conclusion	Final > Initial	Initial > Final

Appendix E

NSCM CONFUSION PATTERNS AS STATE TRANSITION DIAGRAMS
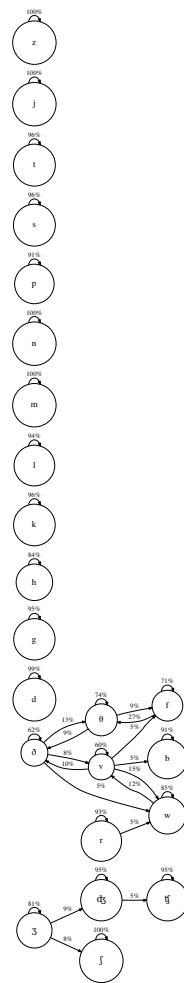
Figure E.1: Anton (RC) CI
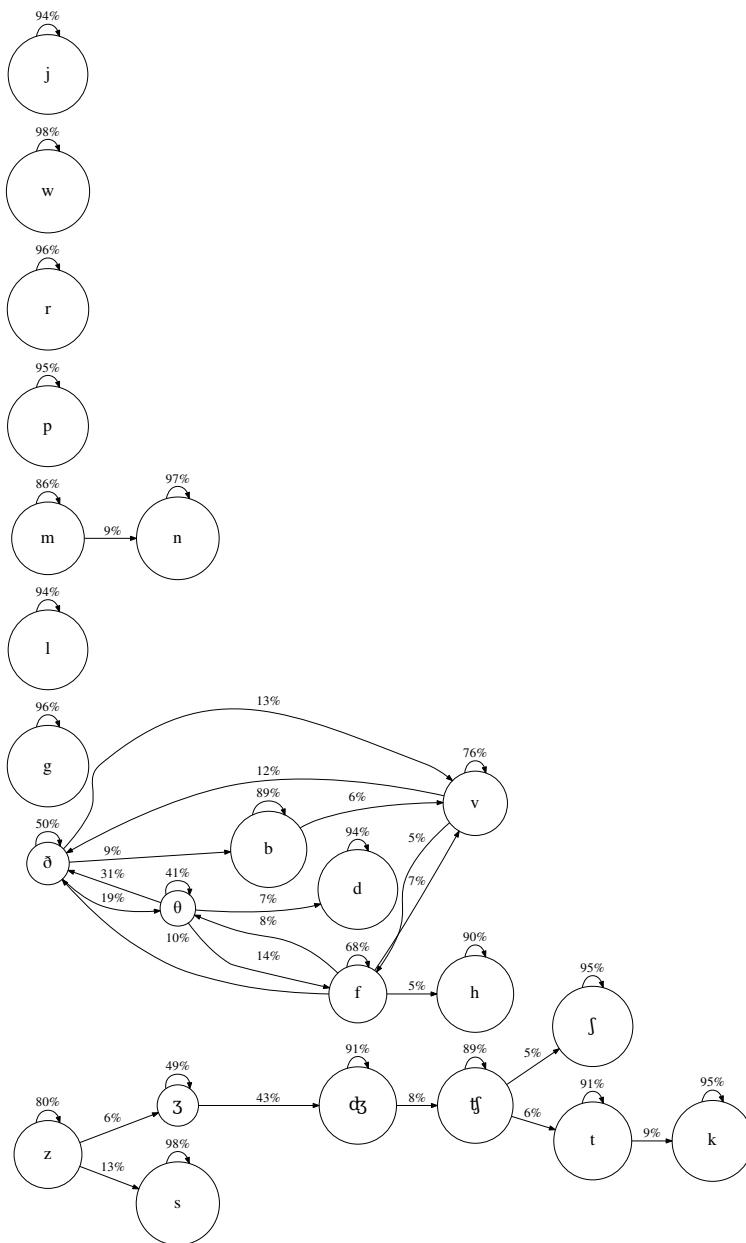
Figure E.2: Bob (RC) CI
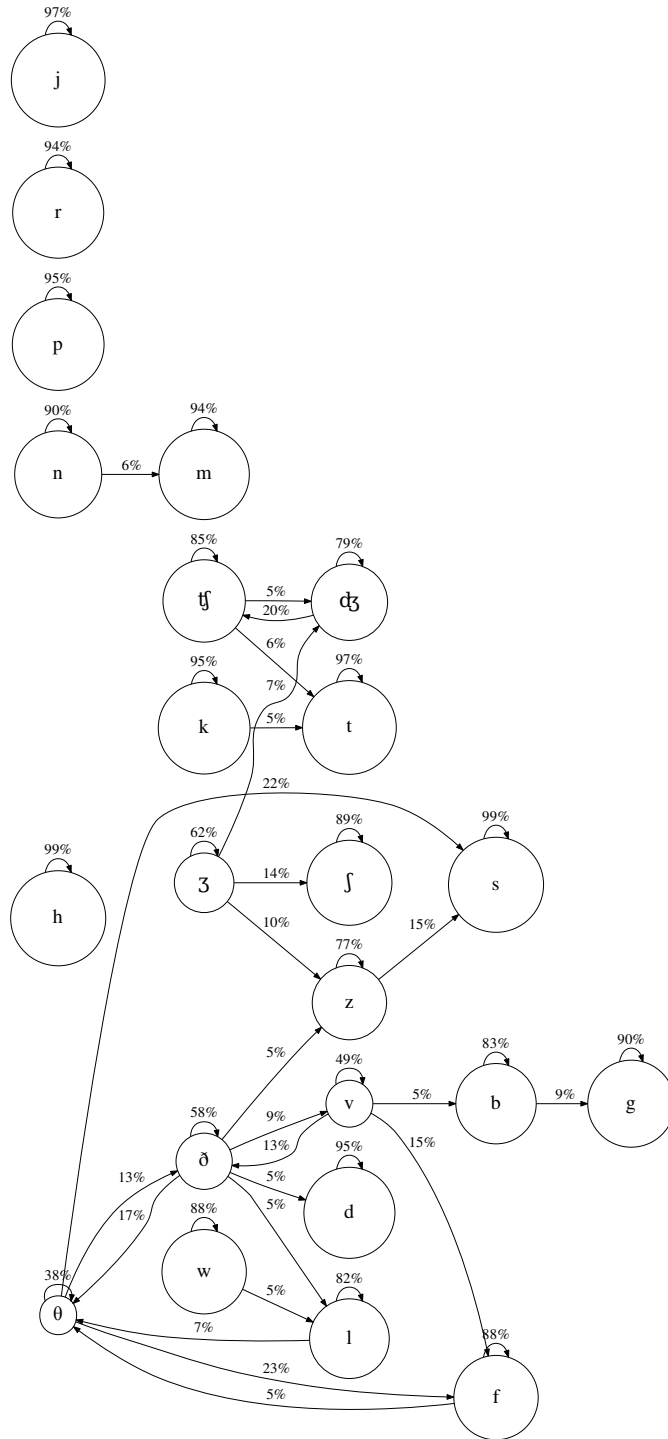
Figure E.3: Carly (RC) CI
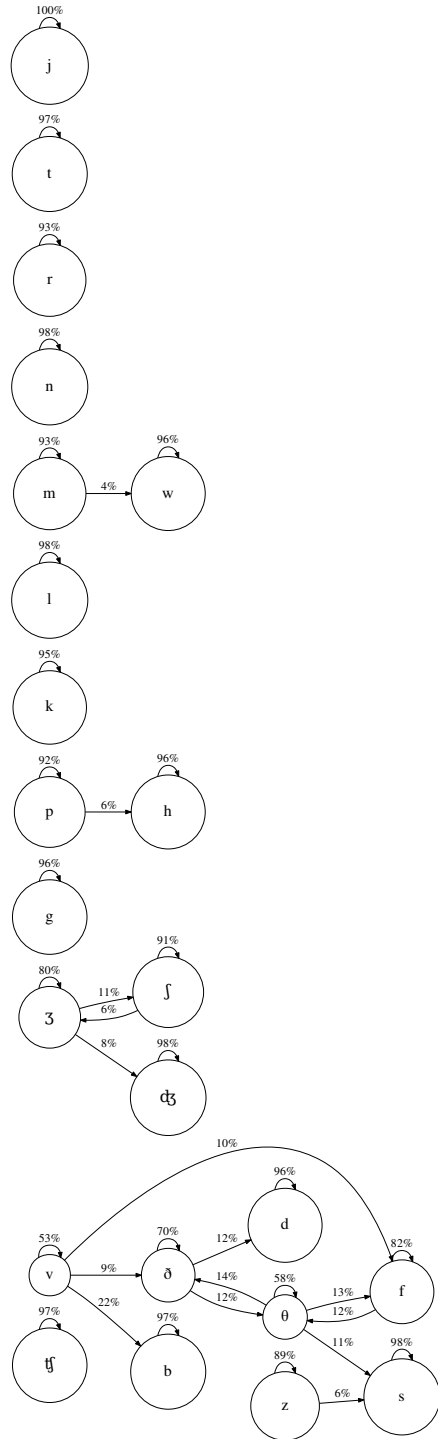
Figure E.4: Evan (RC) CI
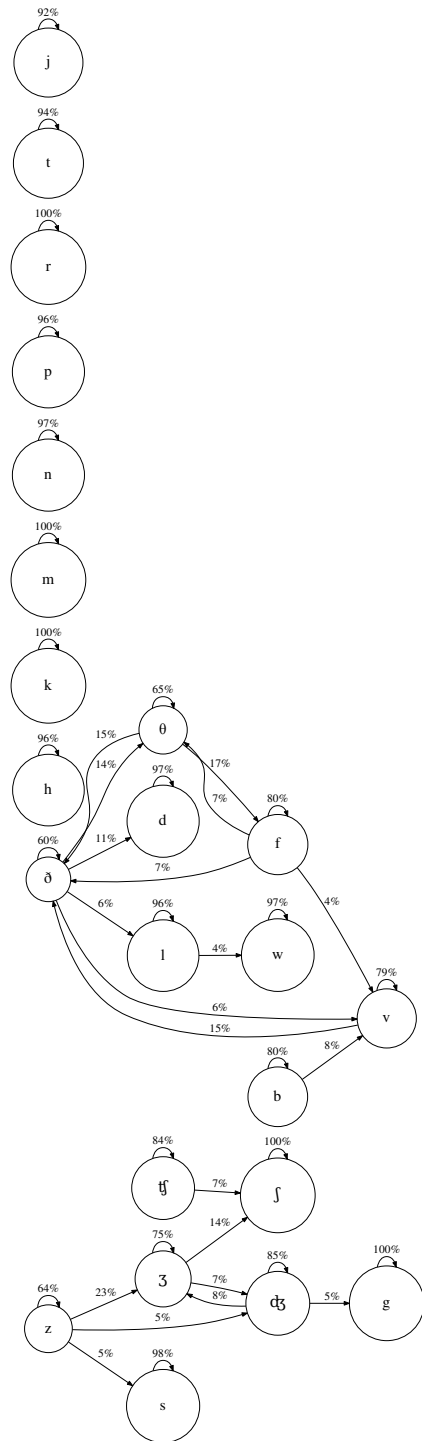
Figure E.5: Joanna (RC) CI
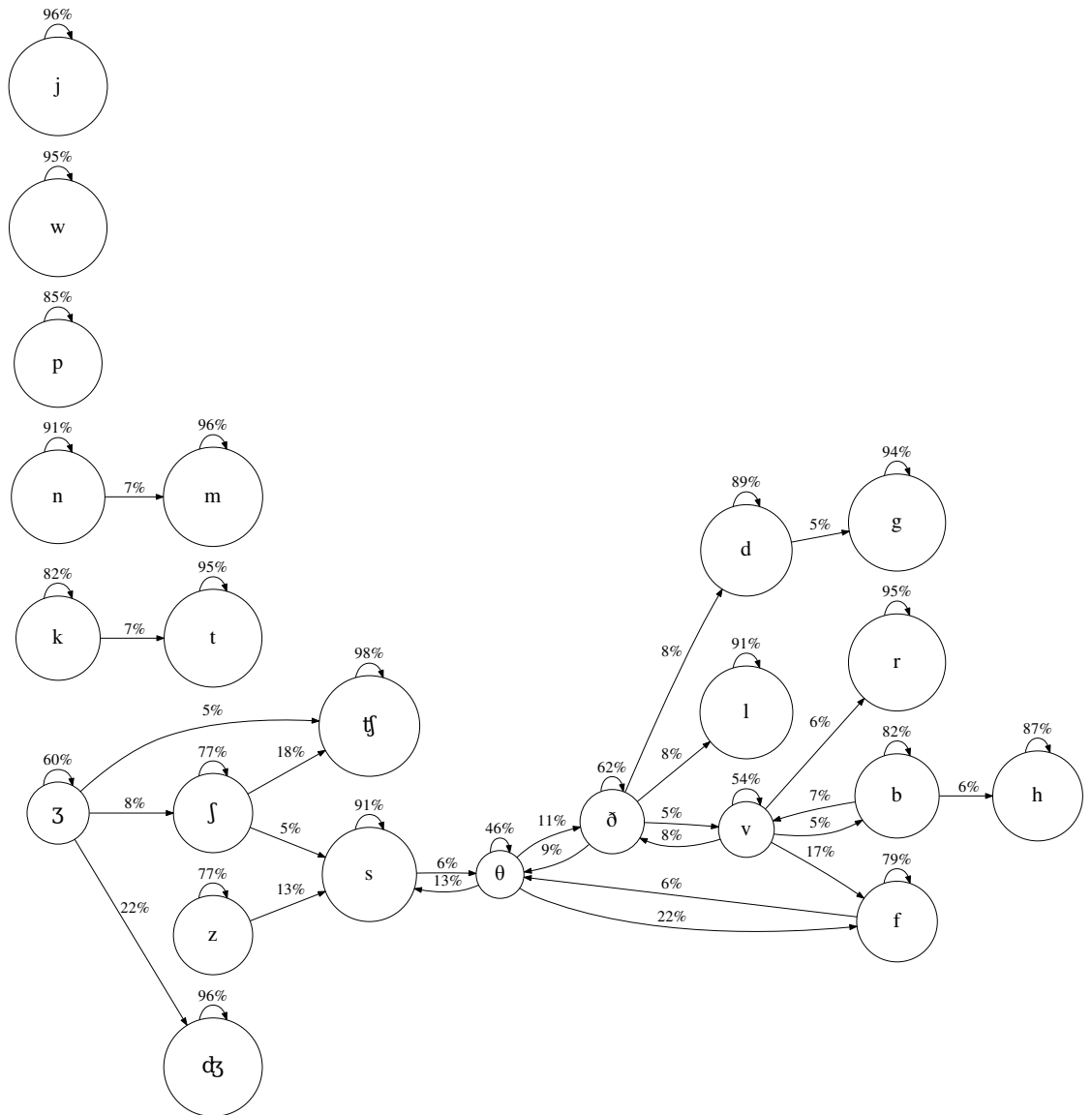
Figure E.6: Miguel (RC) CI
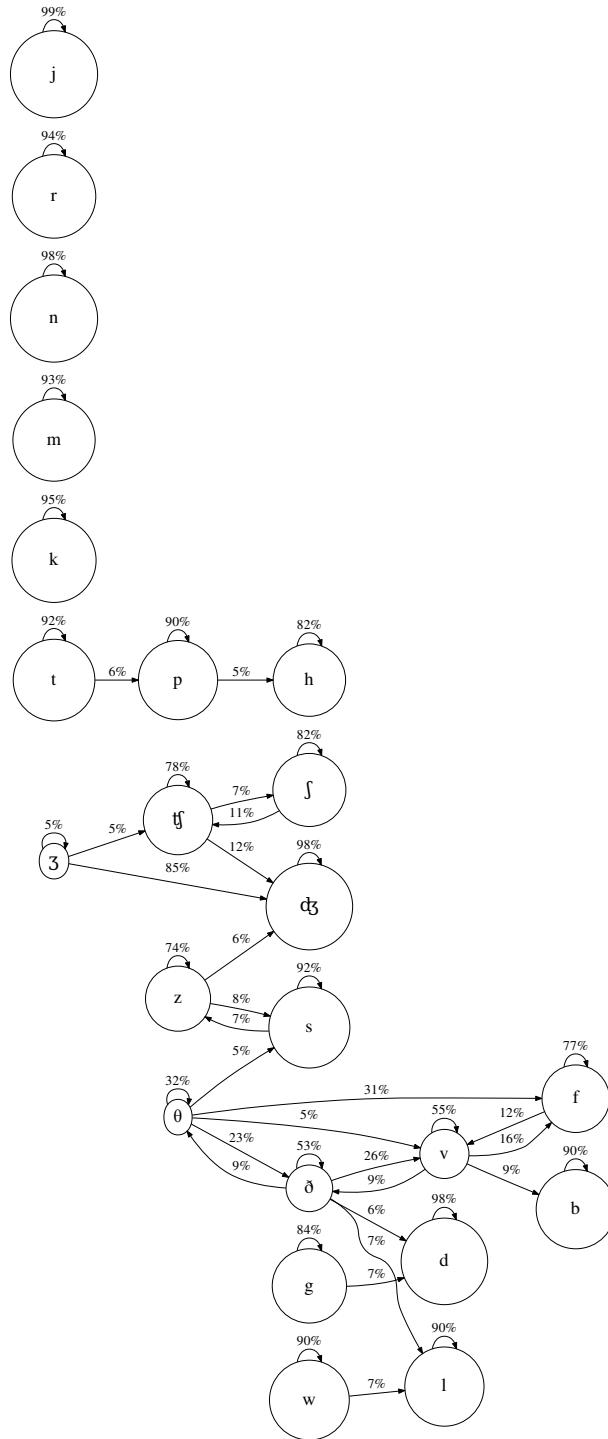
Figure E.7: Alina (RD) CI
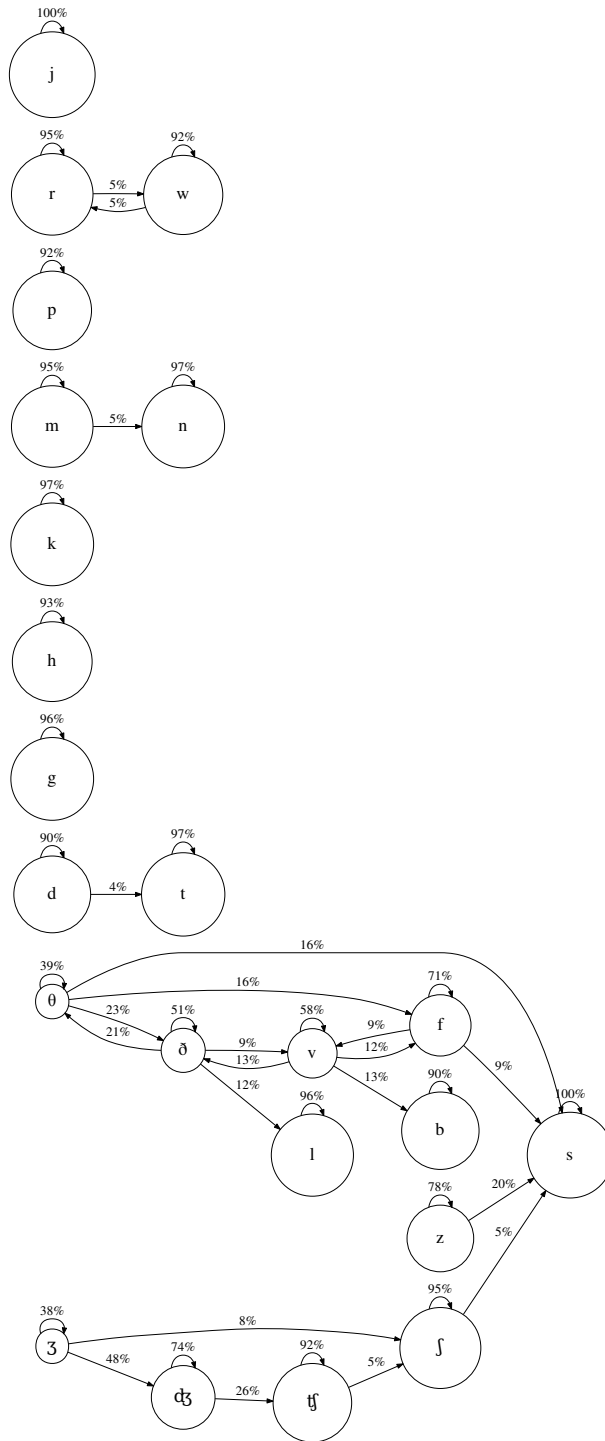
Figure E.8: Angela (RD) CI

101
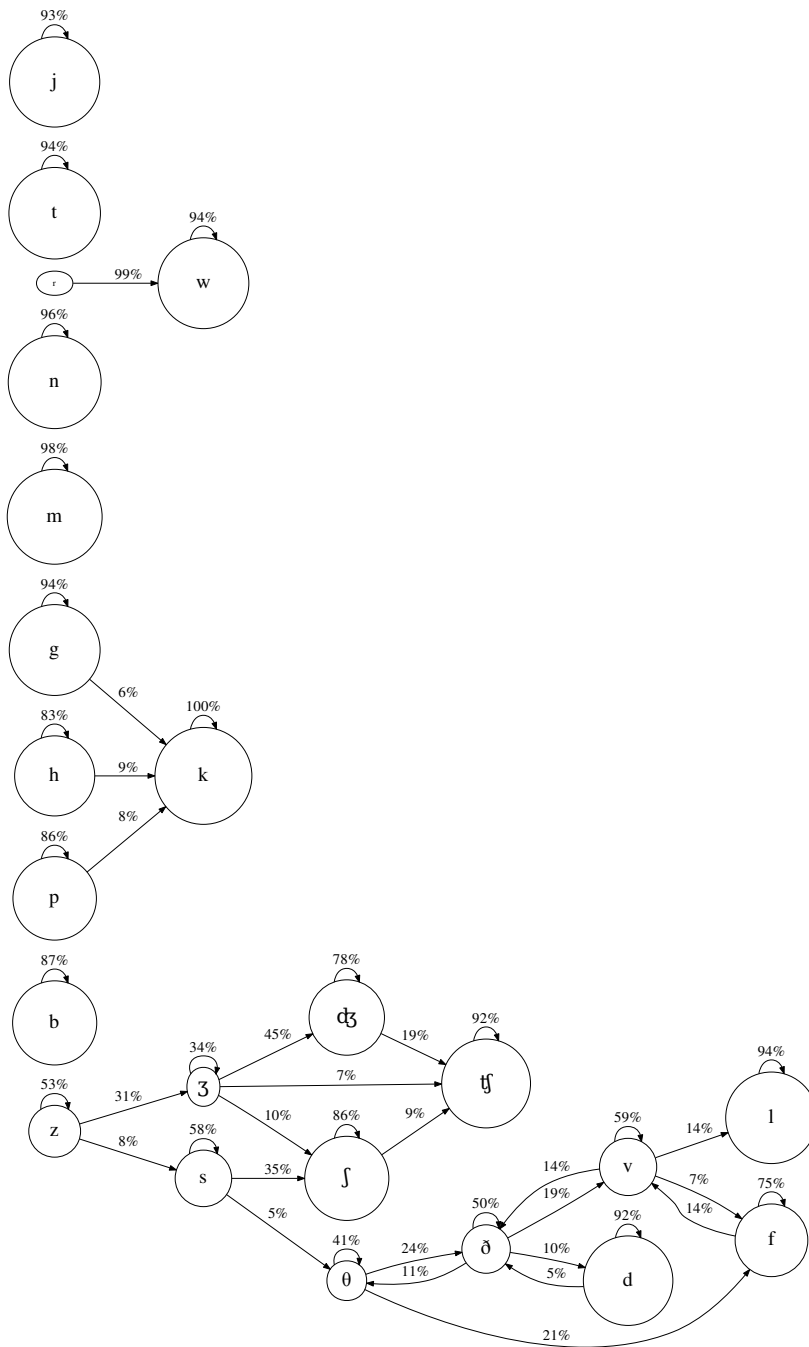
Figure E.9: Edward (RD) CI. Note that Edward could not articulate rhotic phones /r/ or /ɝ/ in his GFTA-2 articulation test. Readers of this graph should ignore any confusion patterns related to these phones.
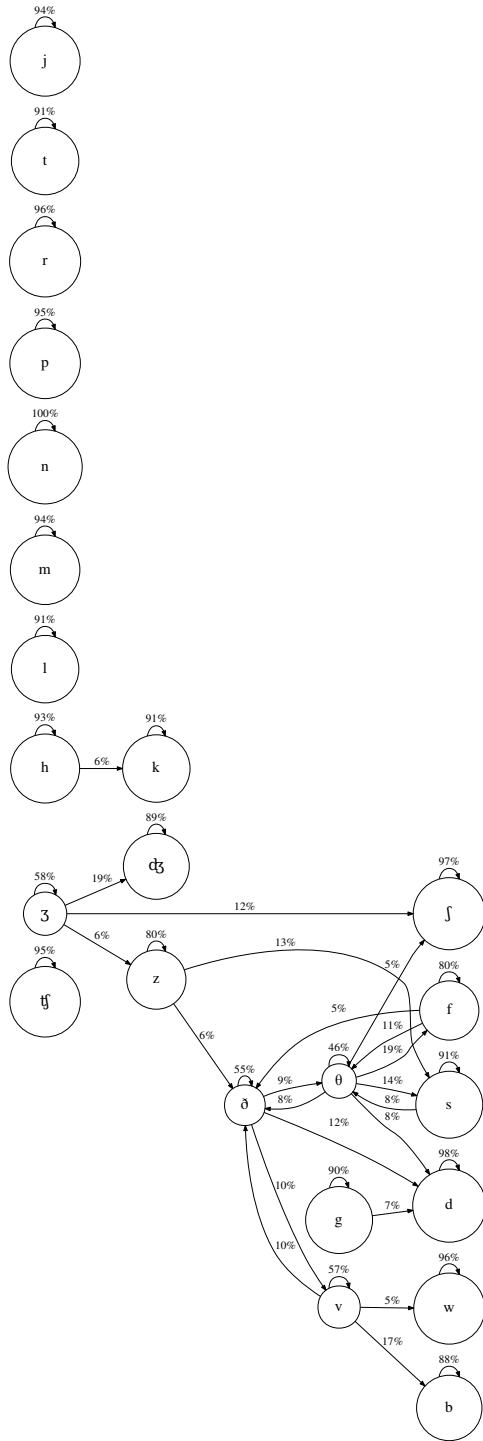
Figure E.10: Latisha (RD) CI
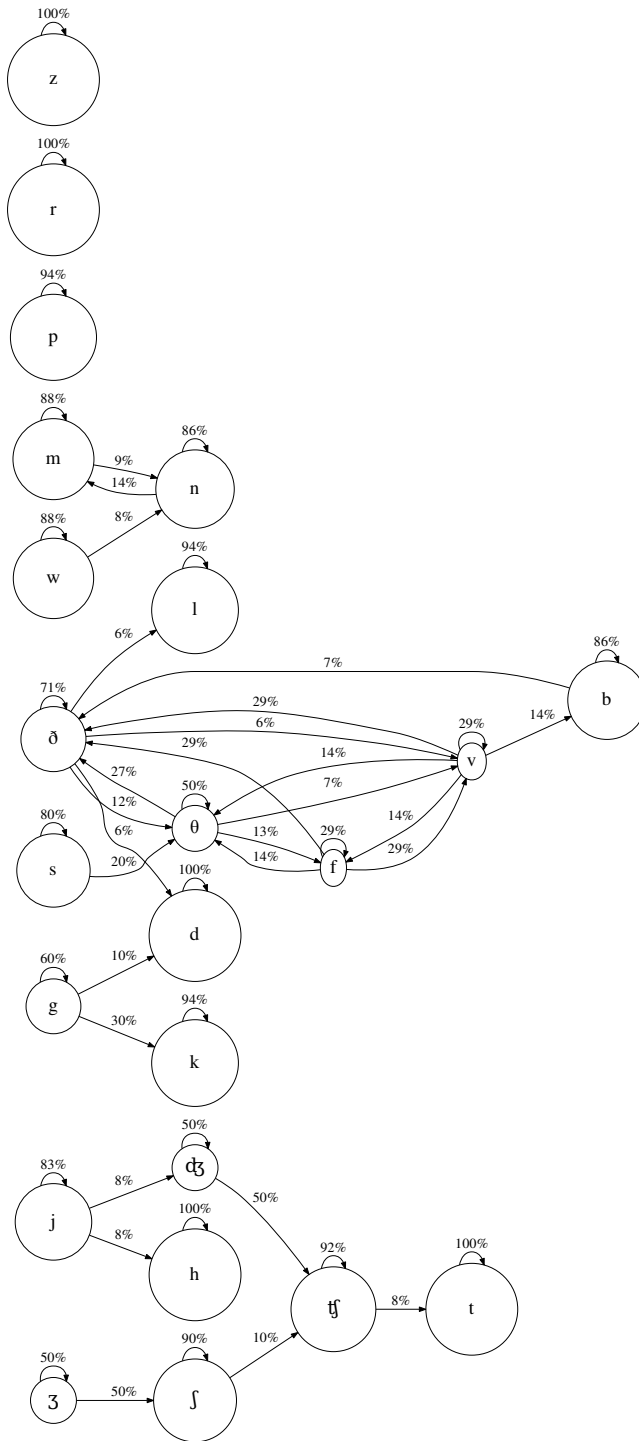
Figure E.11: Laura (RD) CI

104
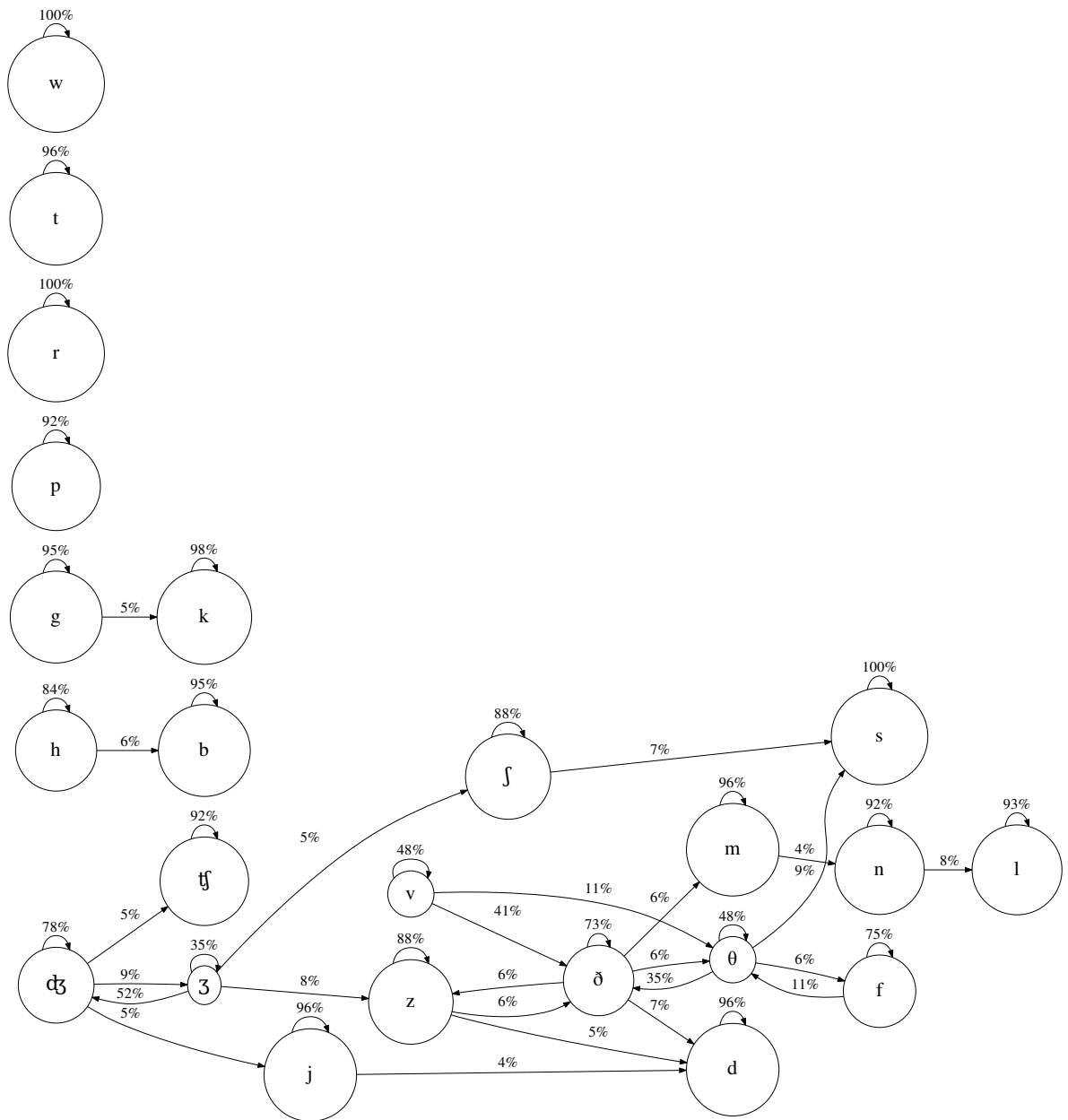
Figure E.12: Norene (RD) CI

105
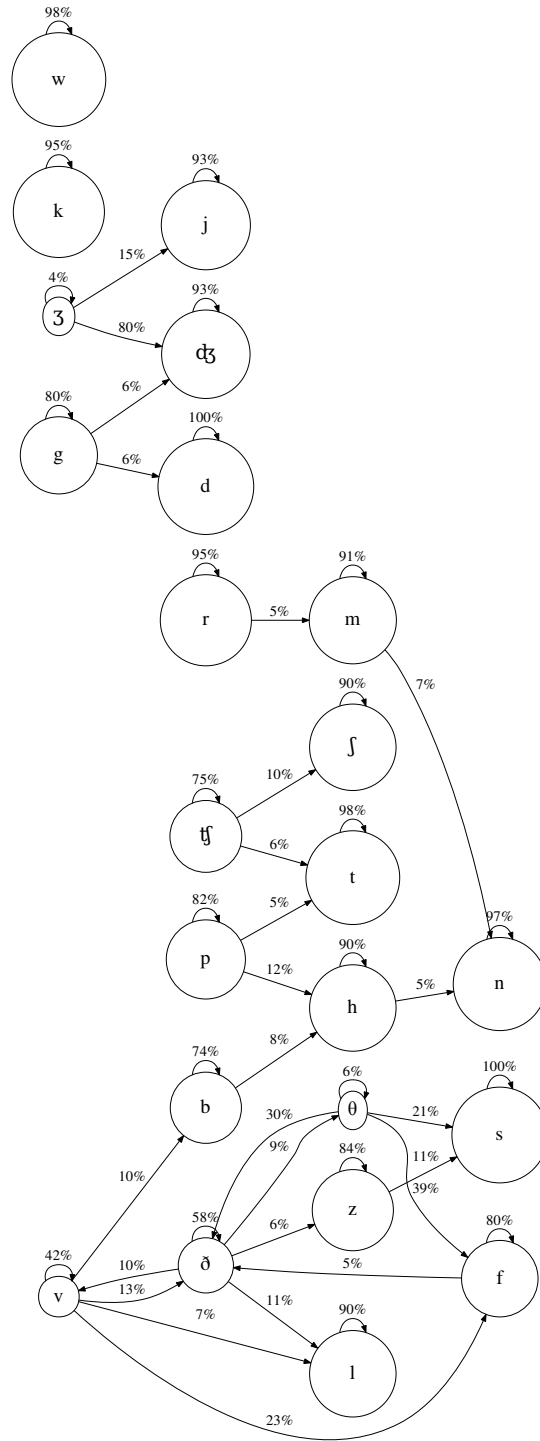
Figure E.13: Shauna (RD) CI
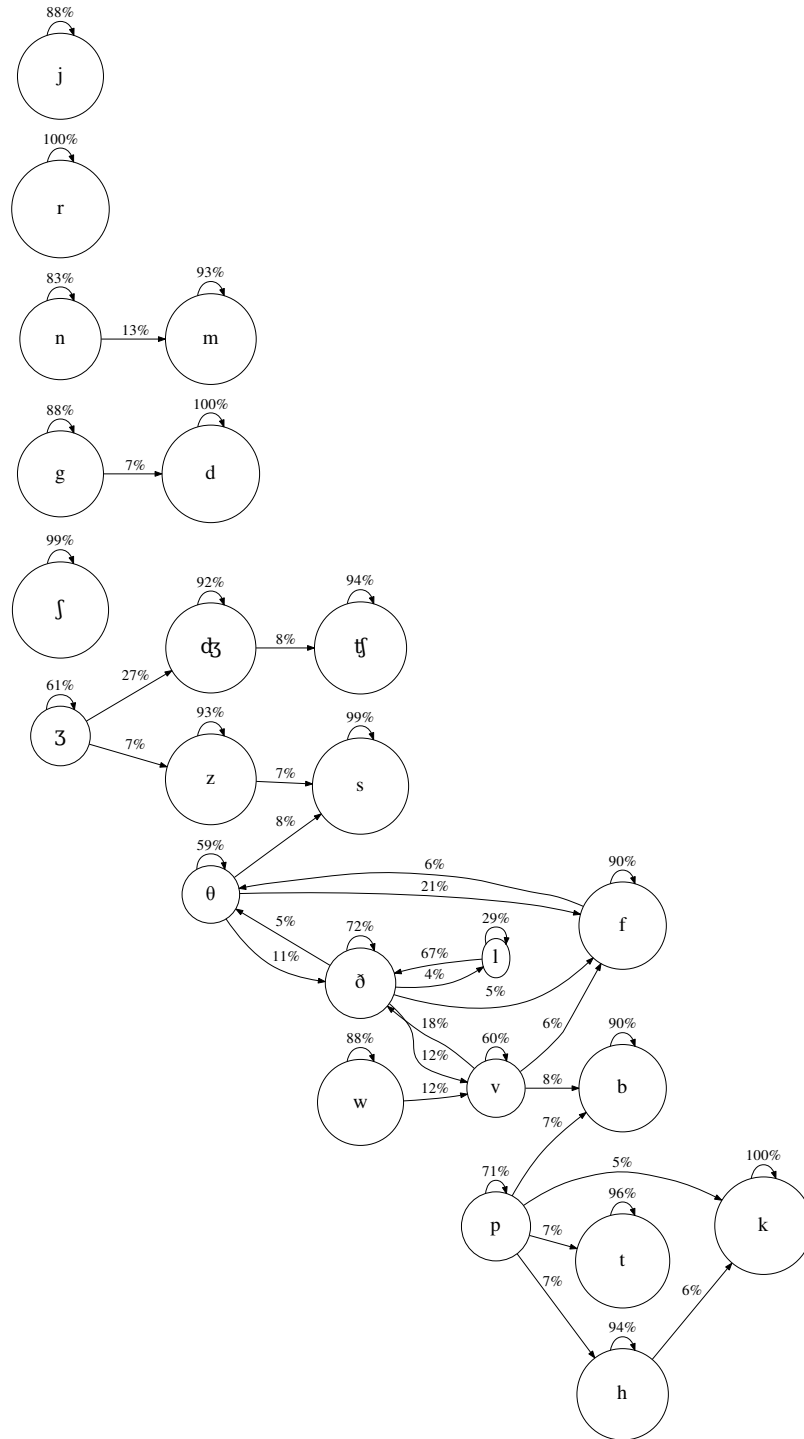
Figure E.14: Teddy (RD) CI
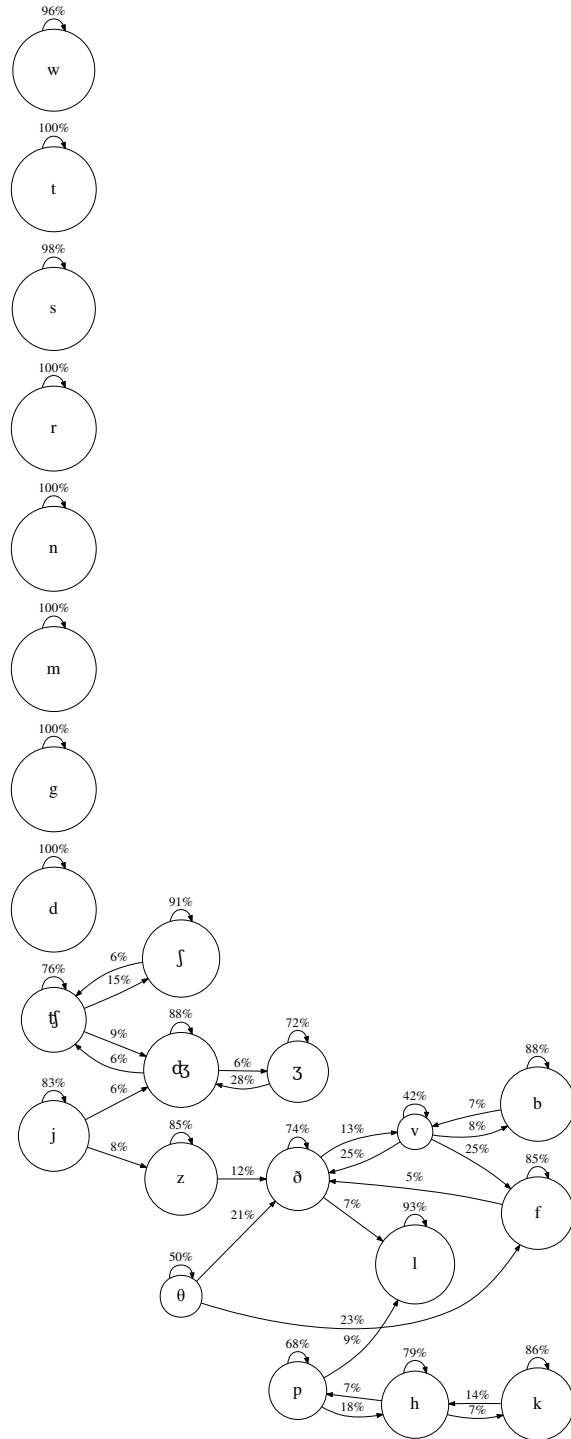
Figure E.15: Tony (RD) CI
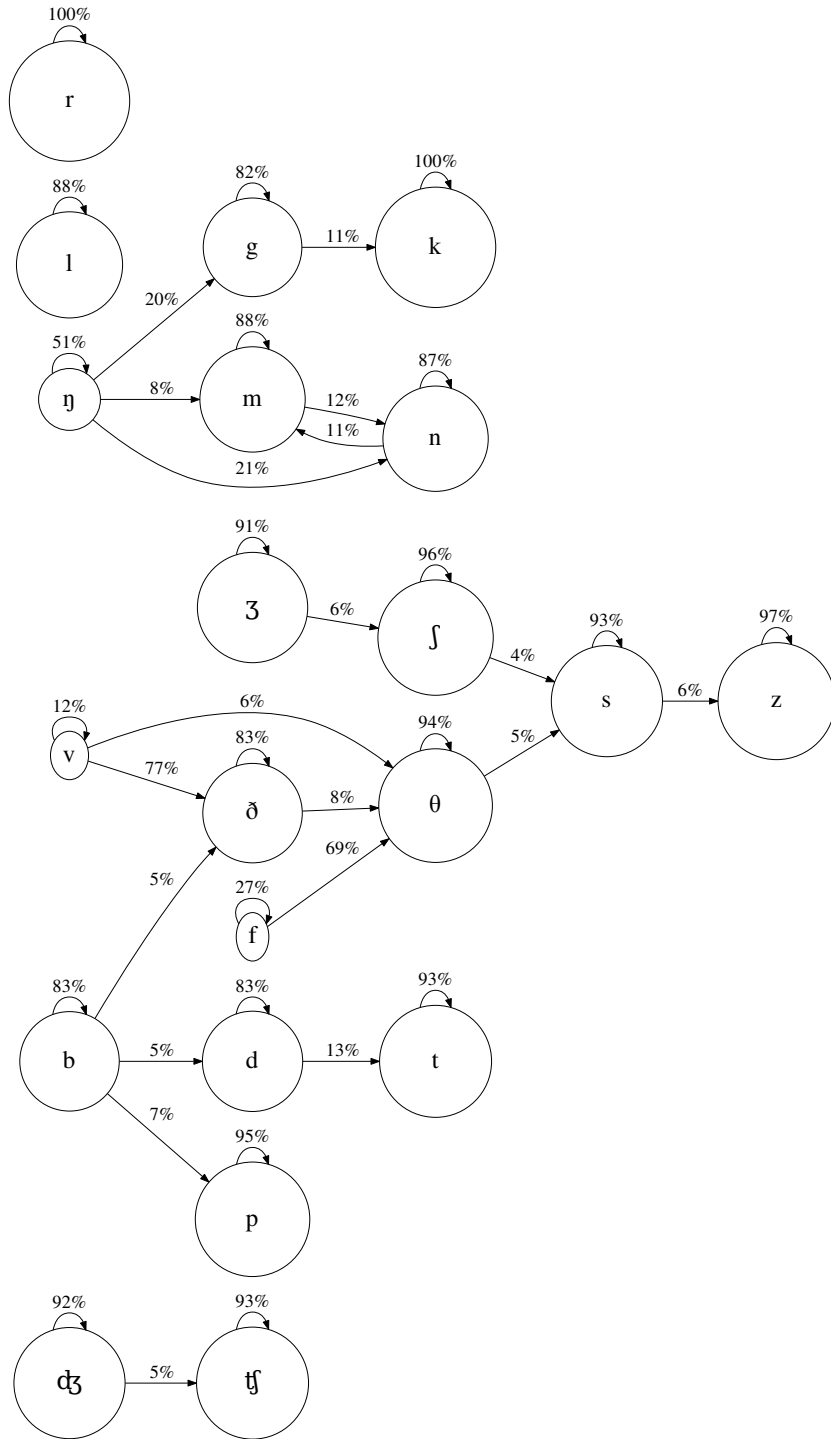
Figure E.16: Anton (RC) CF

109
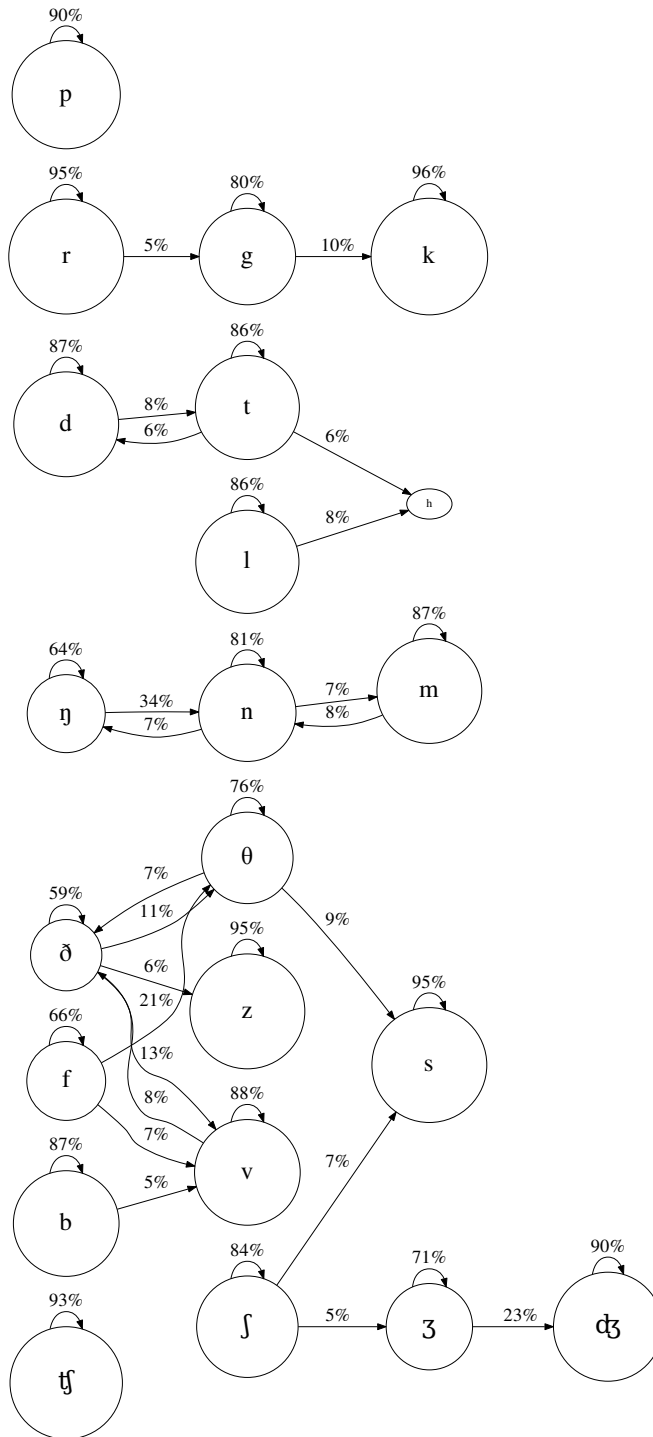
Figure E.17: Bob (RC) CF
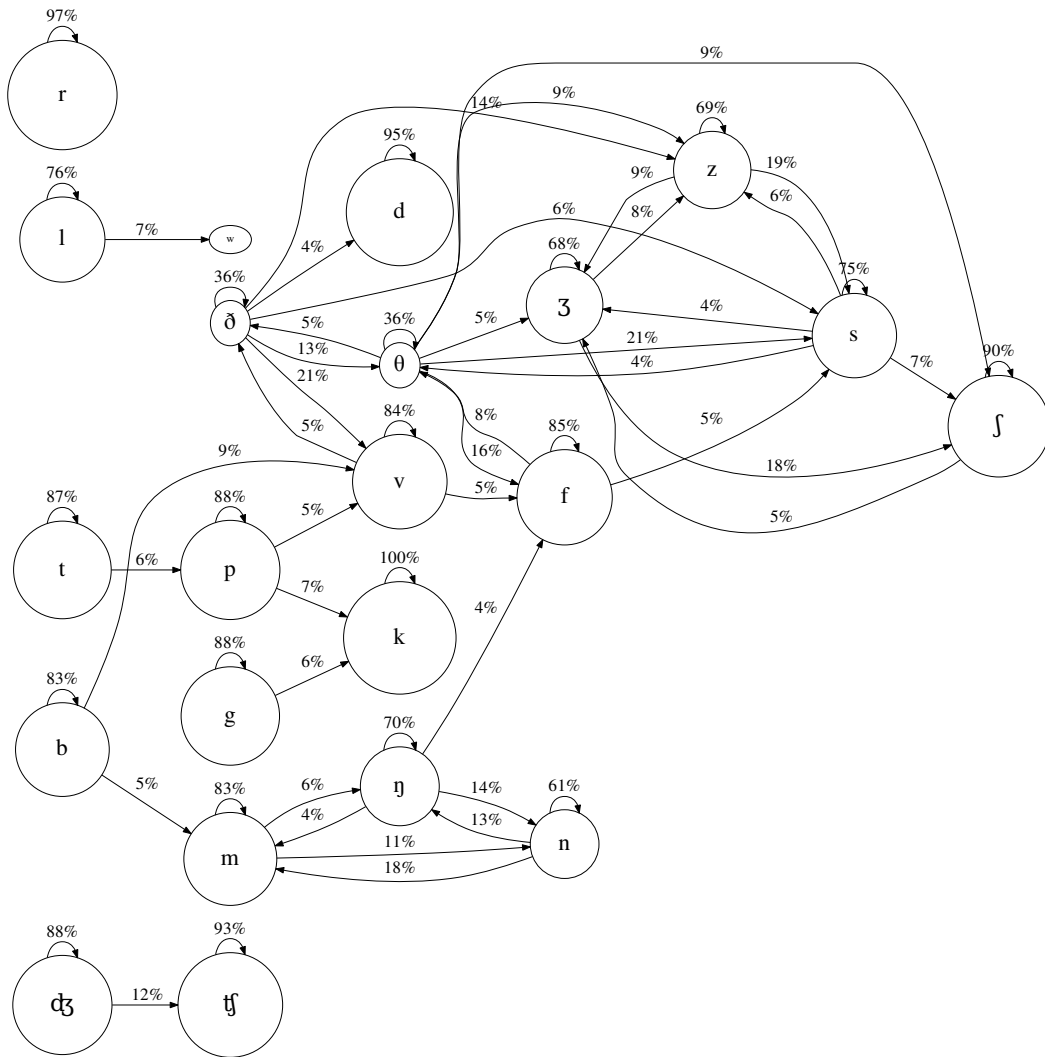
Figure E.18: Carly (RC) CF
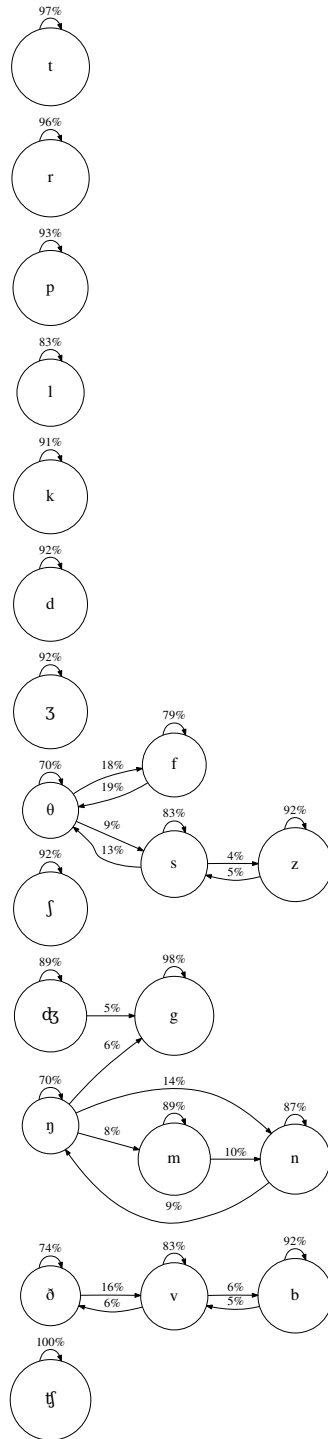
Figure E.19: Evan (RC) CF
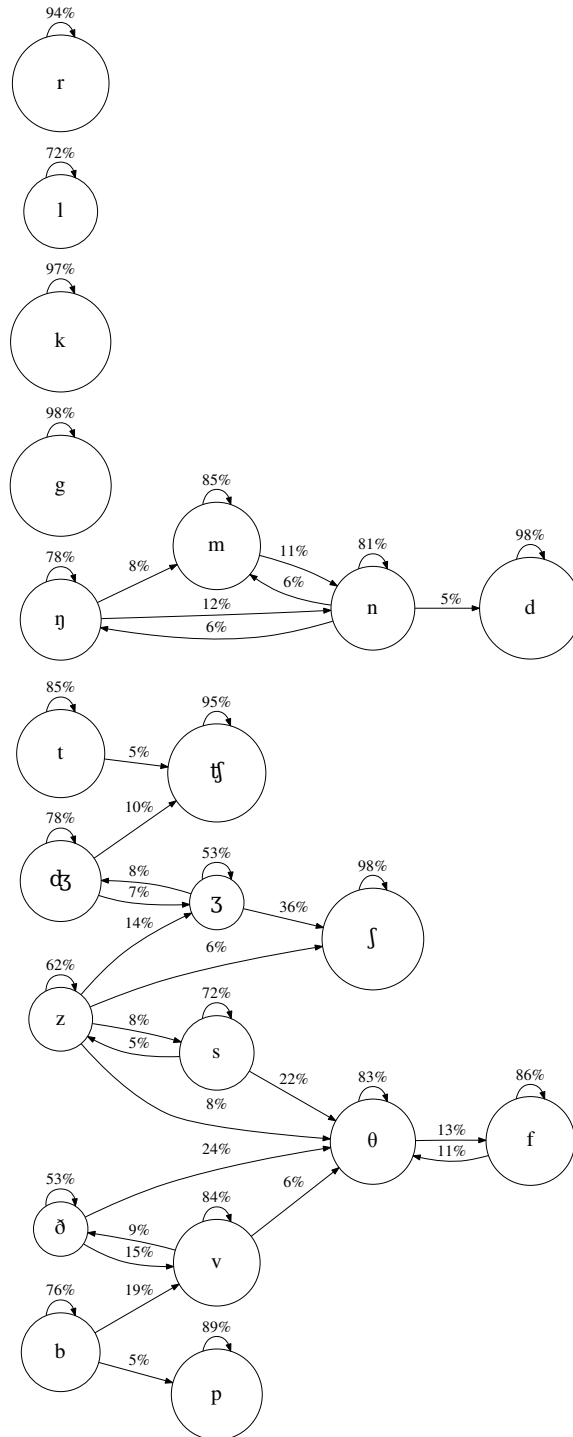
Figure E.20: Joanna (RC) CF

113
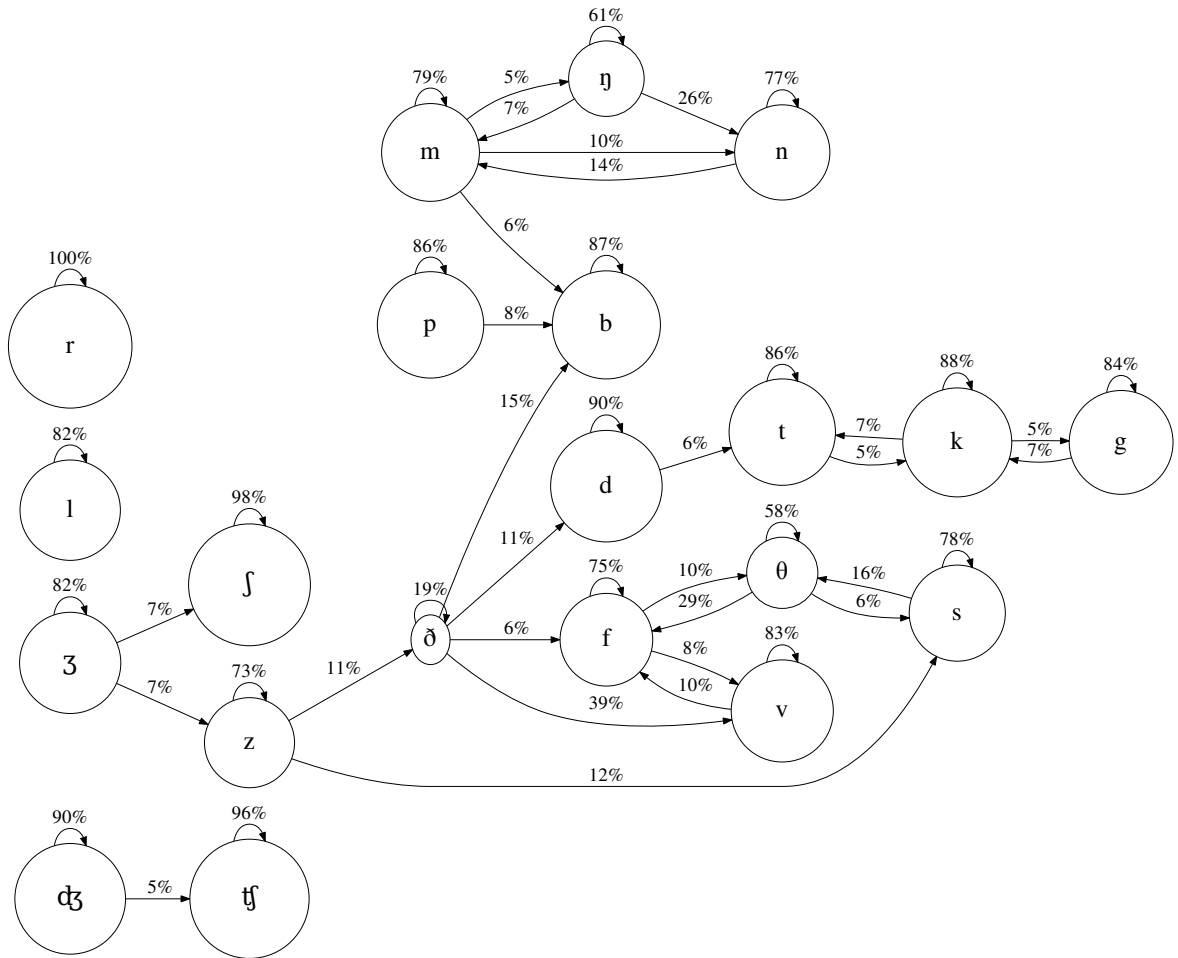
Figure E.21: Miguel (RC) CF
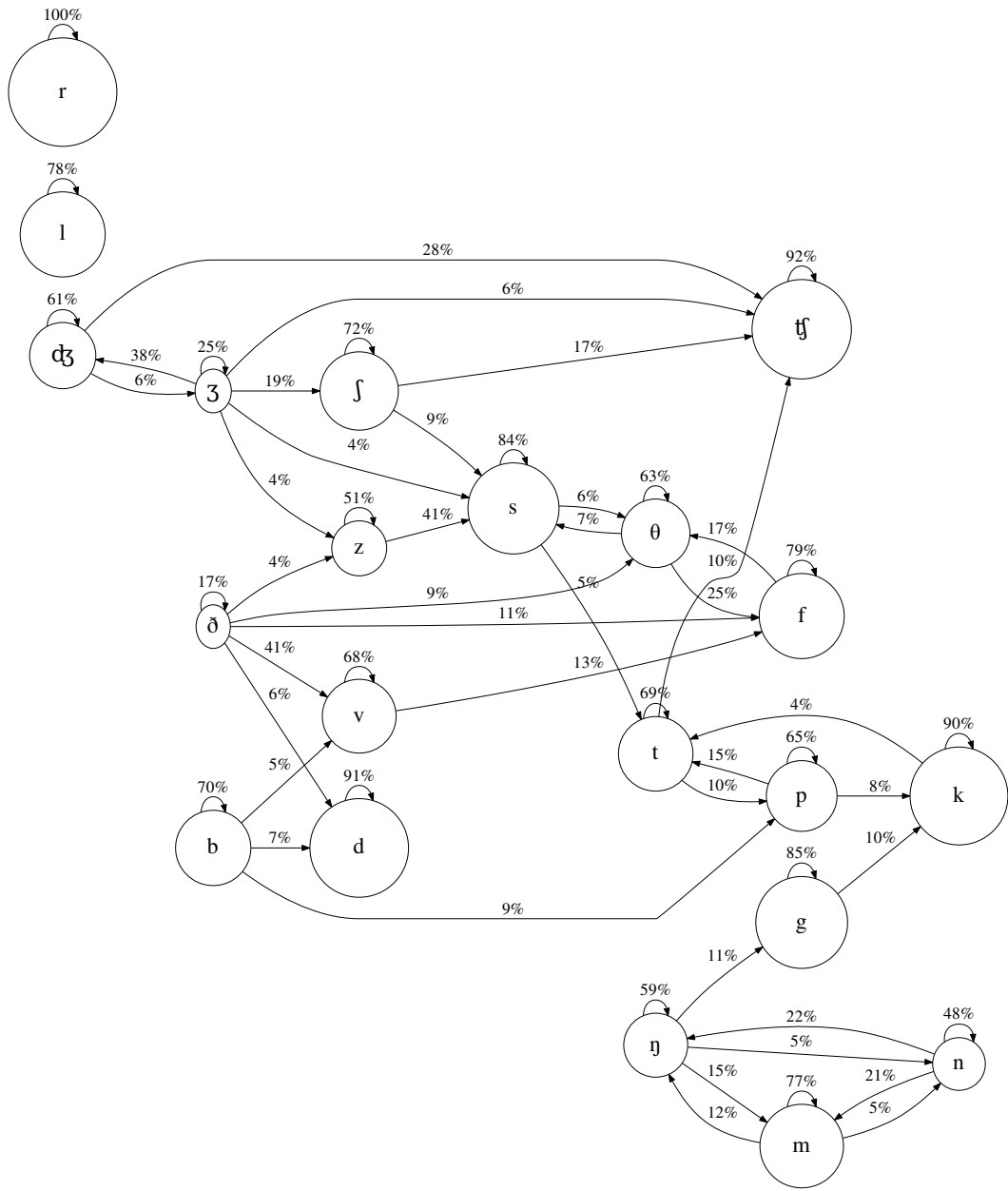
Figure E.22: Alina (RD) CF
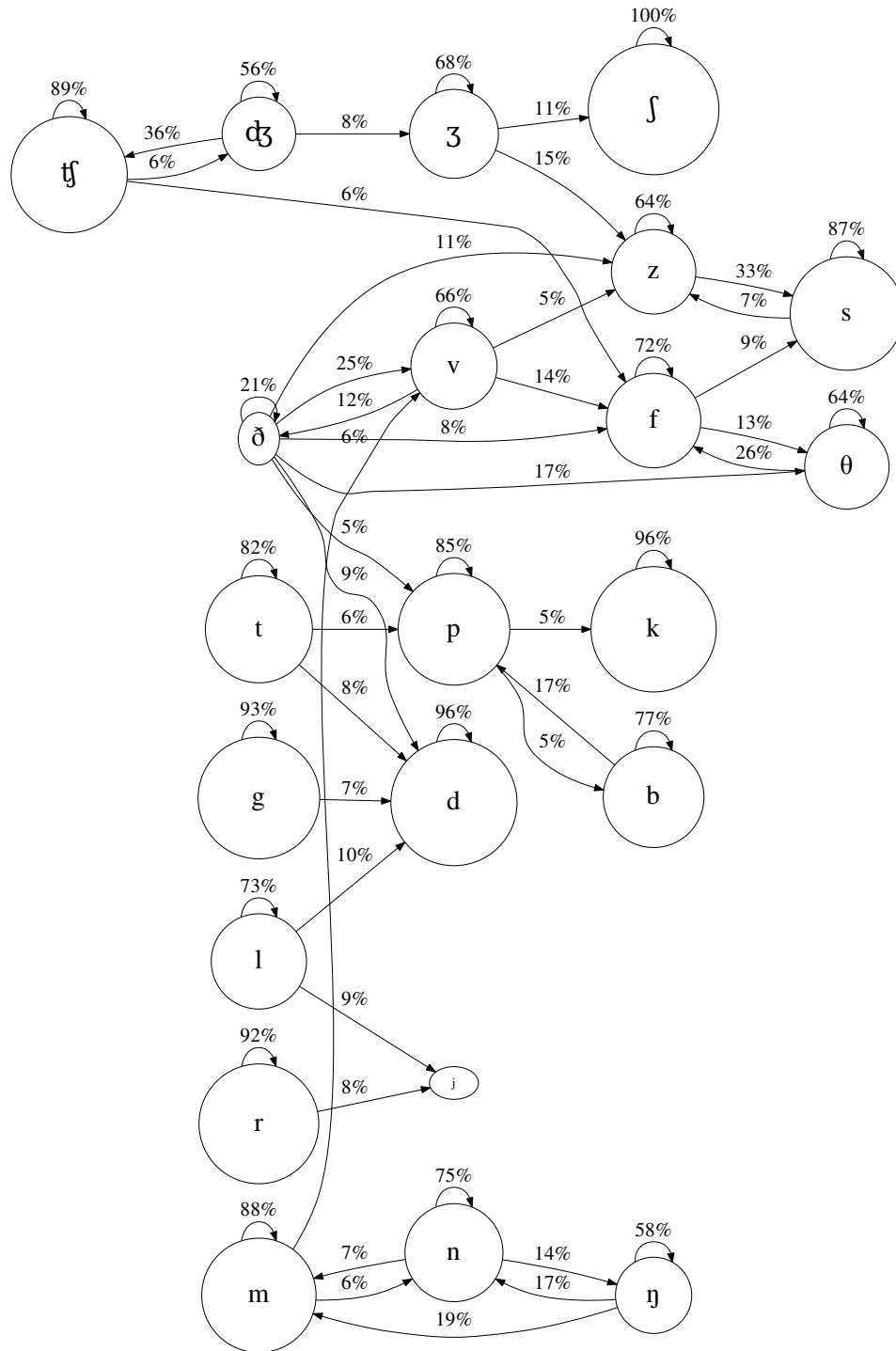
Figure E.23: Angela (RD) CF

116
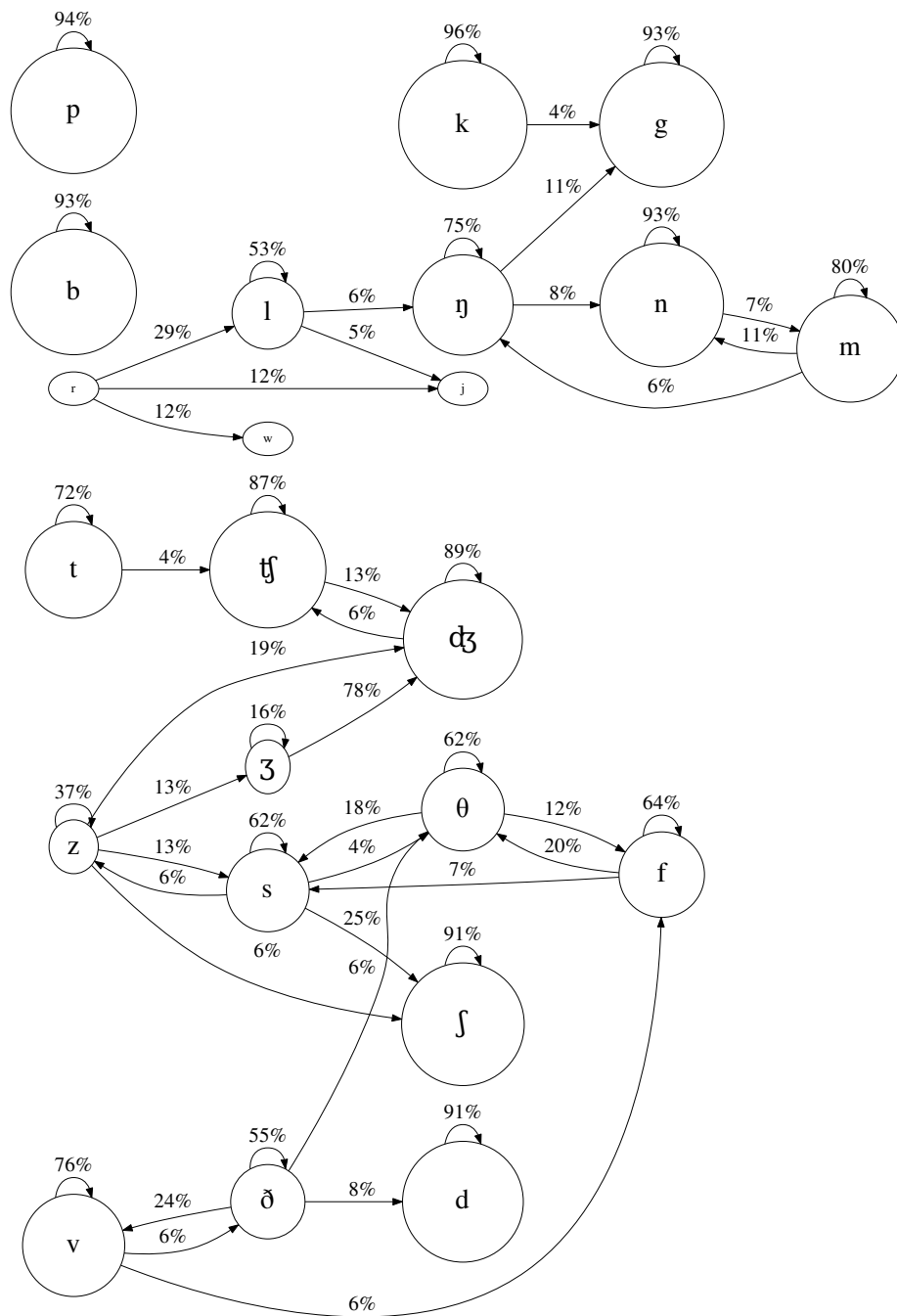
Figure E.24: Edward (RD) CF. Note that Edward could not articulate rhotic phones /r/ or /ɝ/ in his GFTA-2 articulation test. Readers of this graph should ignore any confusion patterns related to these phones.
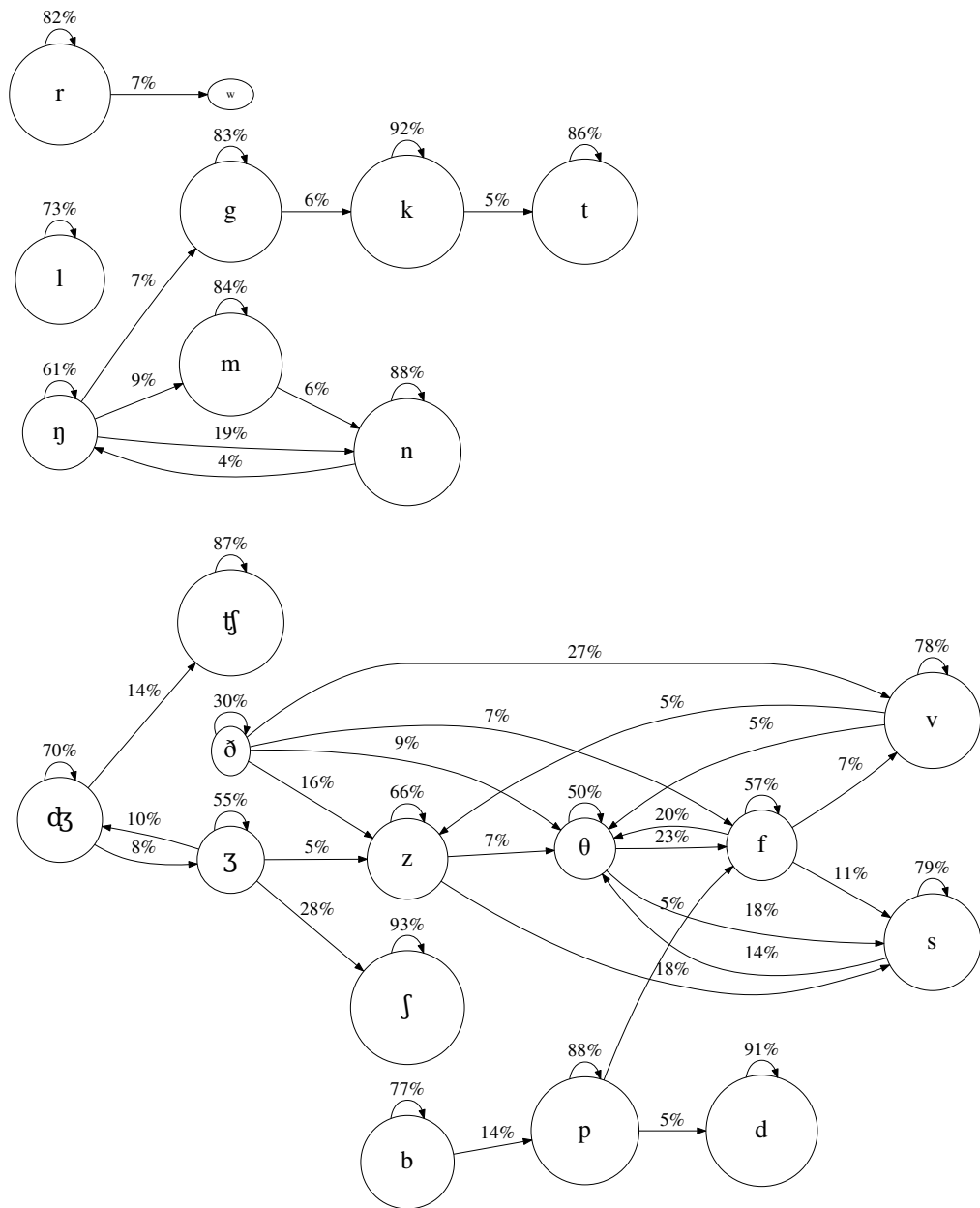
Figure E.25: Latisha (RD) CF

Figure E.26: Laura (RD) CF

119

Figure E.27: Norene (RD) CF

120
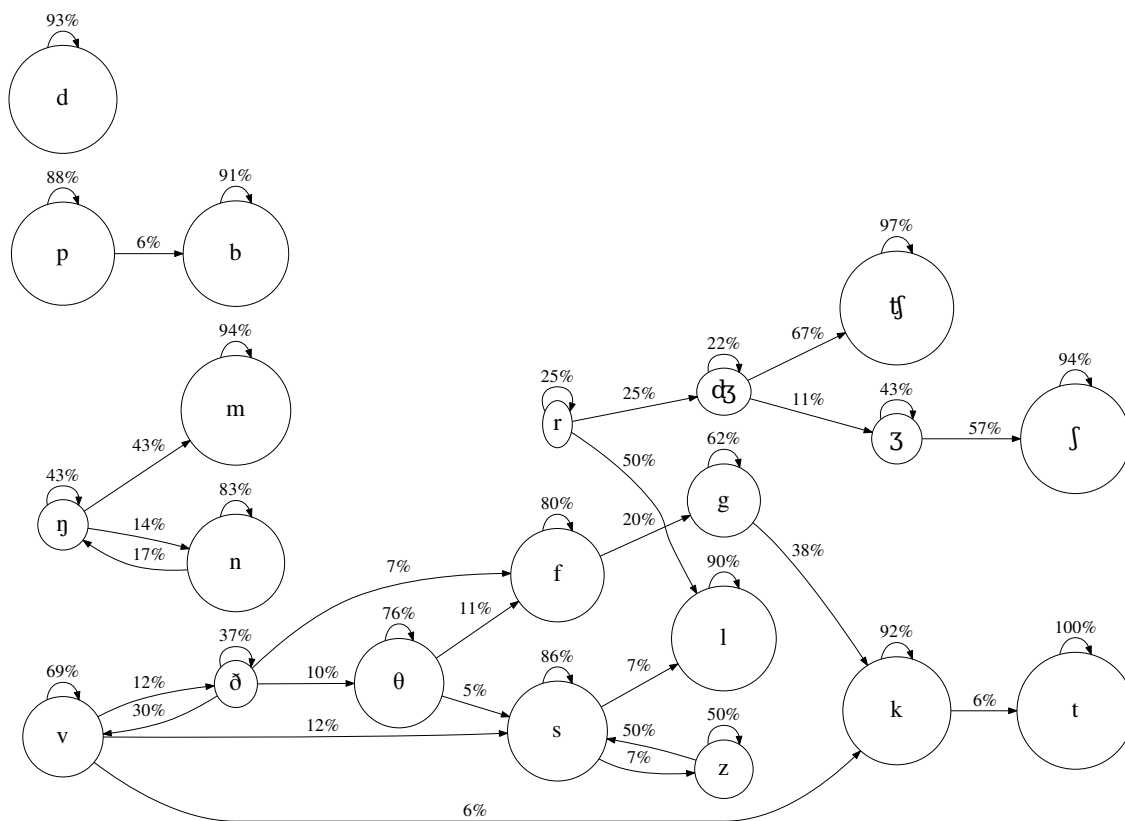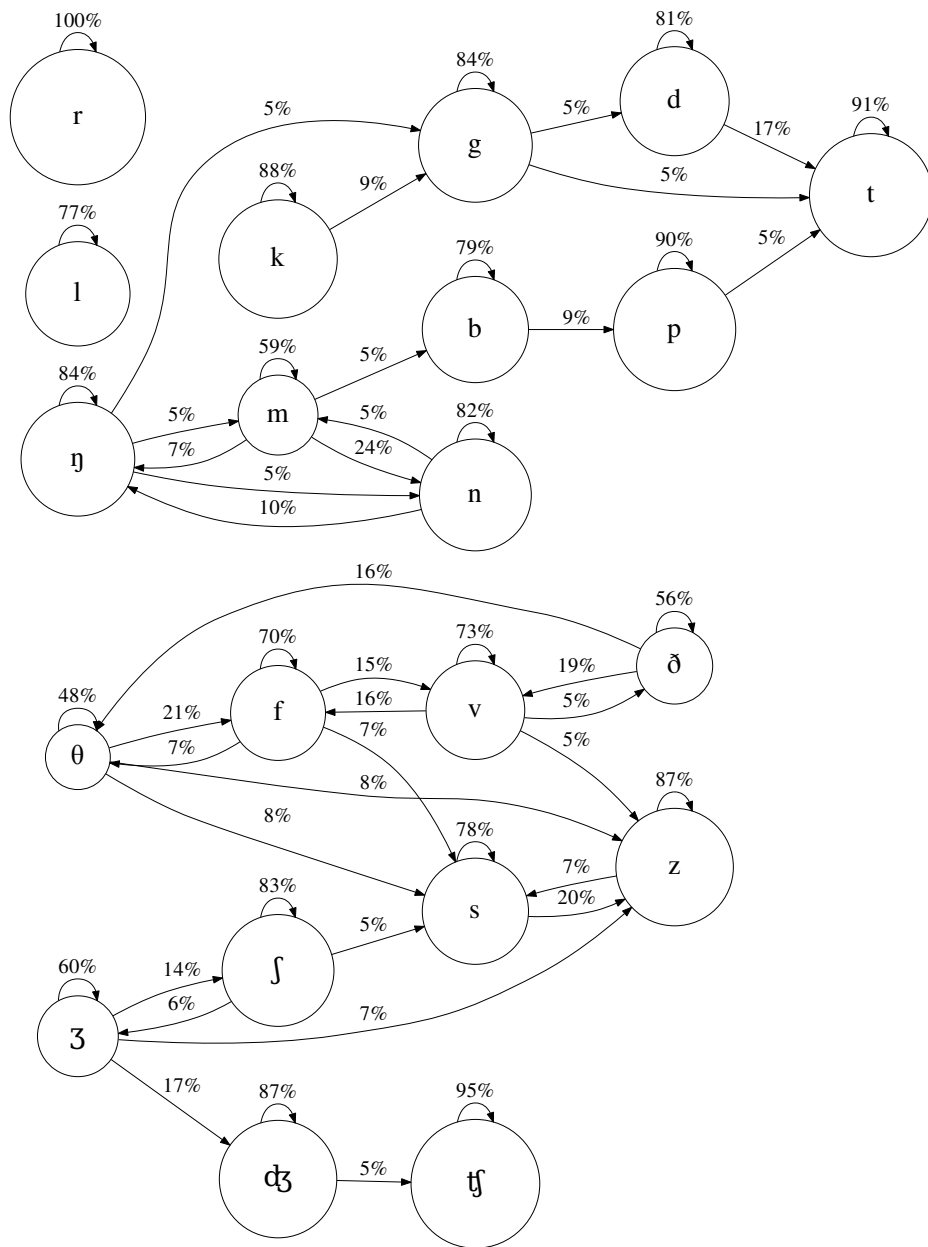
Figure E.28: Shauna (RD) CF

121
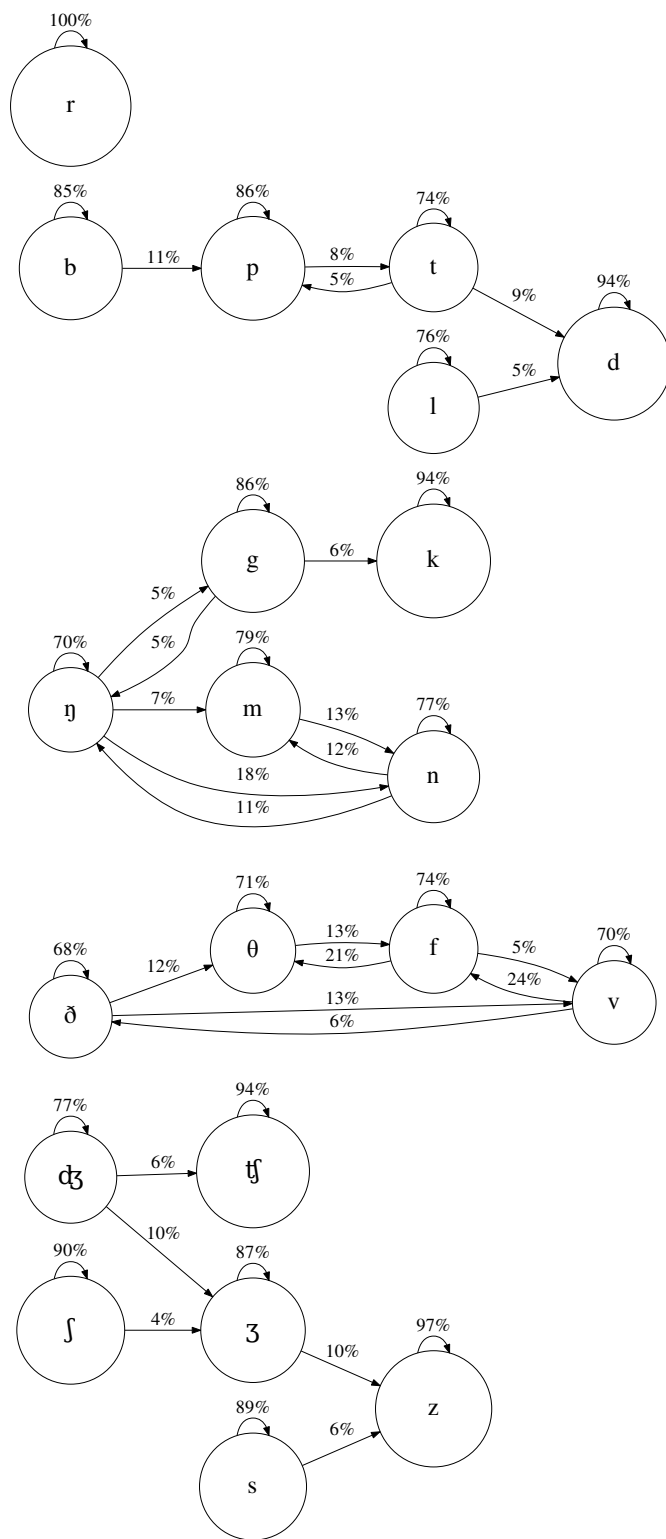
Figure E.29: Teddy (RD) CF

122
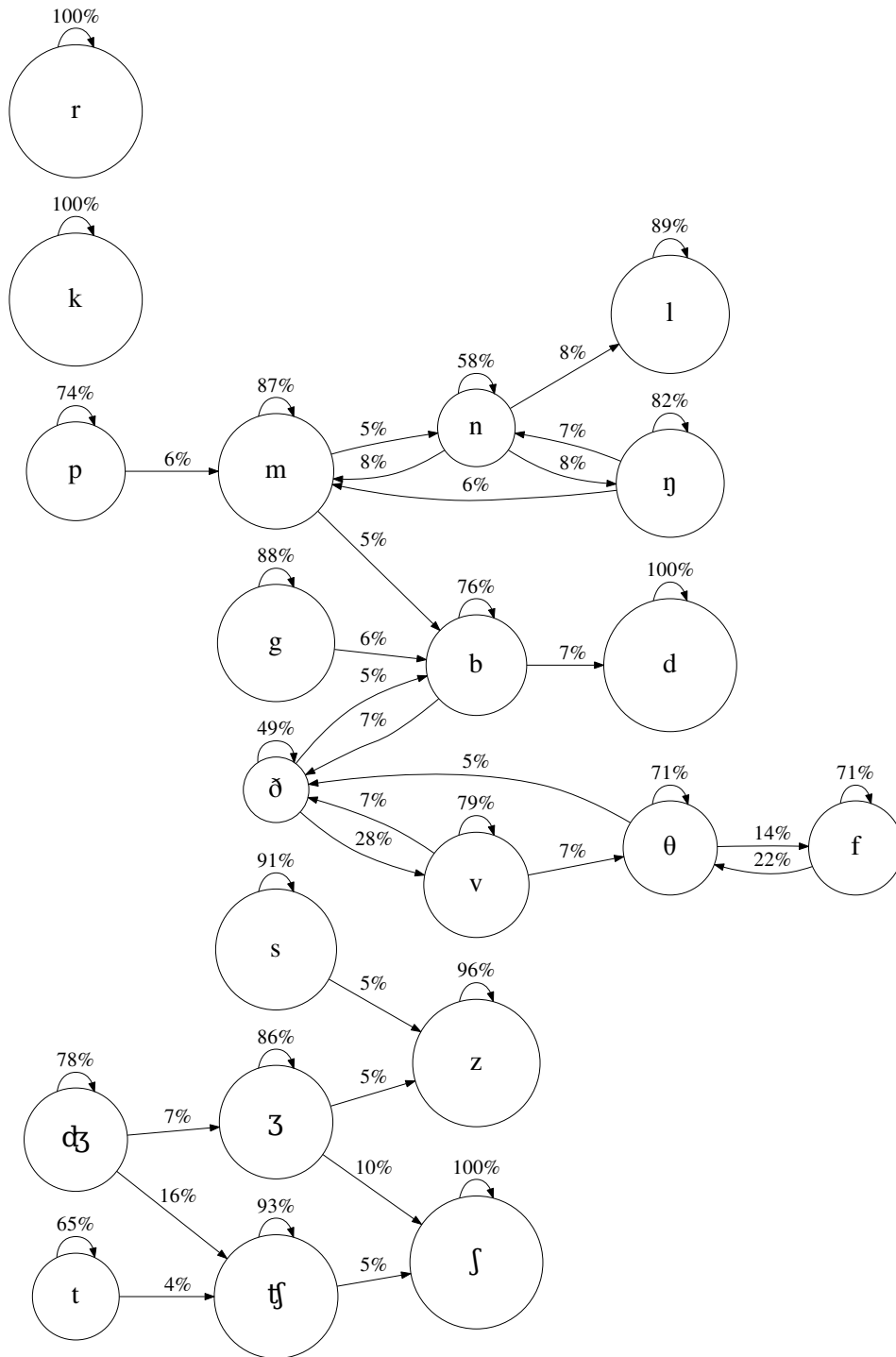
Figure E.30: Tony (RD) CF

123
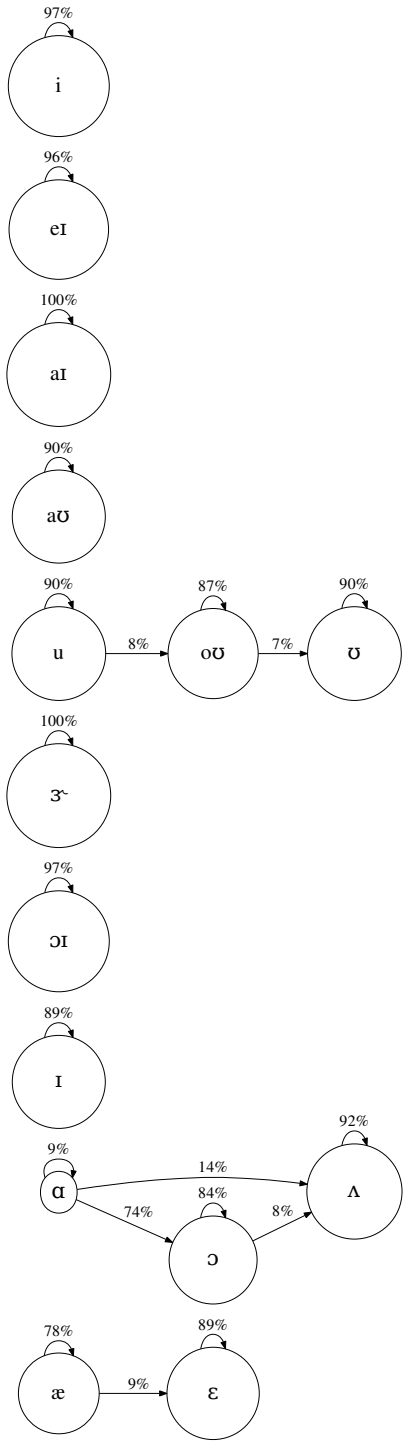
Figure E.31: Anton (RC) VI
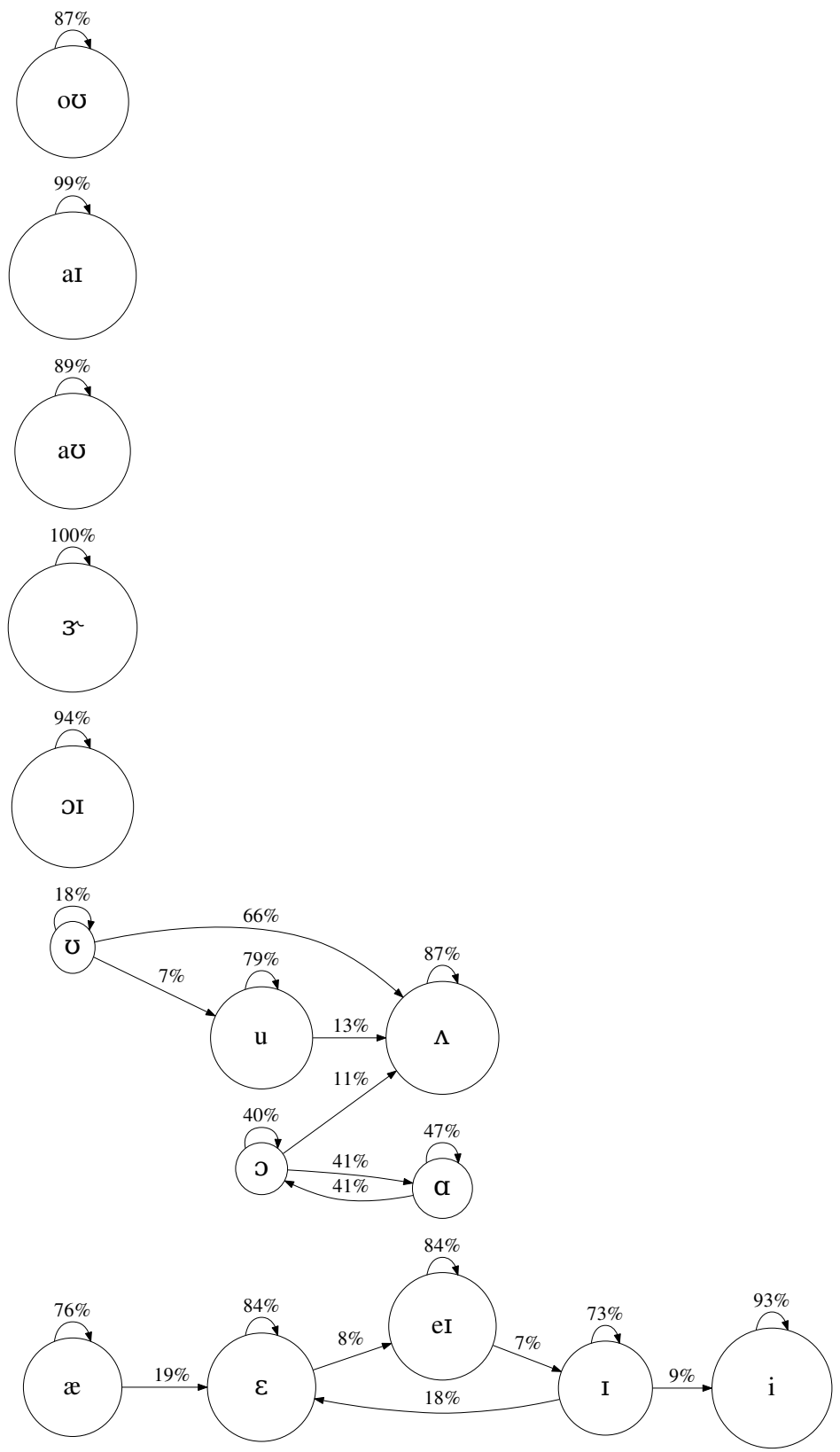
Figure E.32: Bob (RC) VI

125
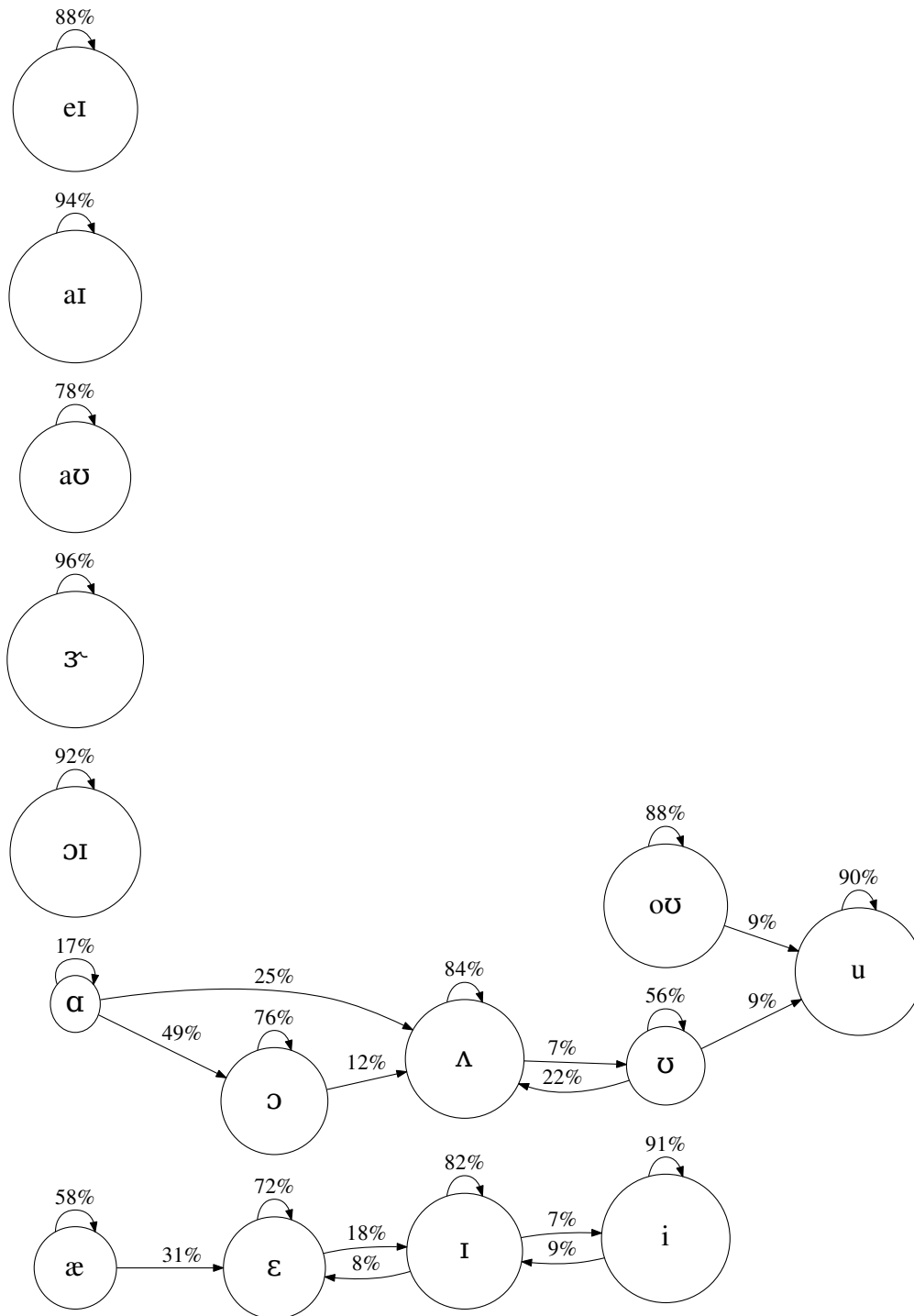
Figure E.33: Carly (RC) VI

126
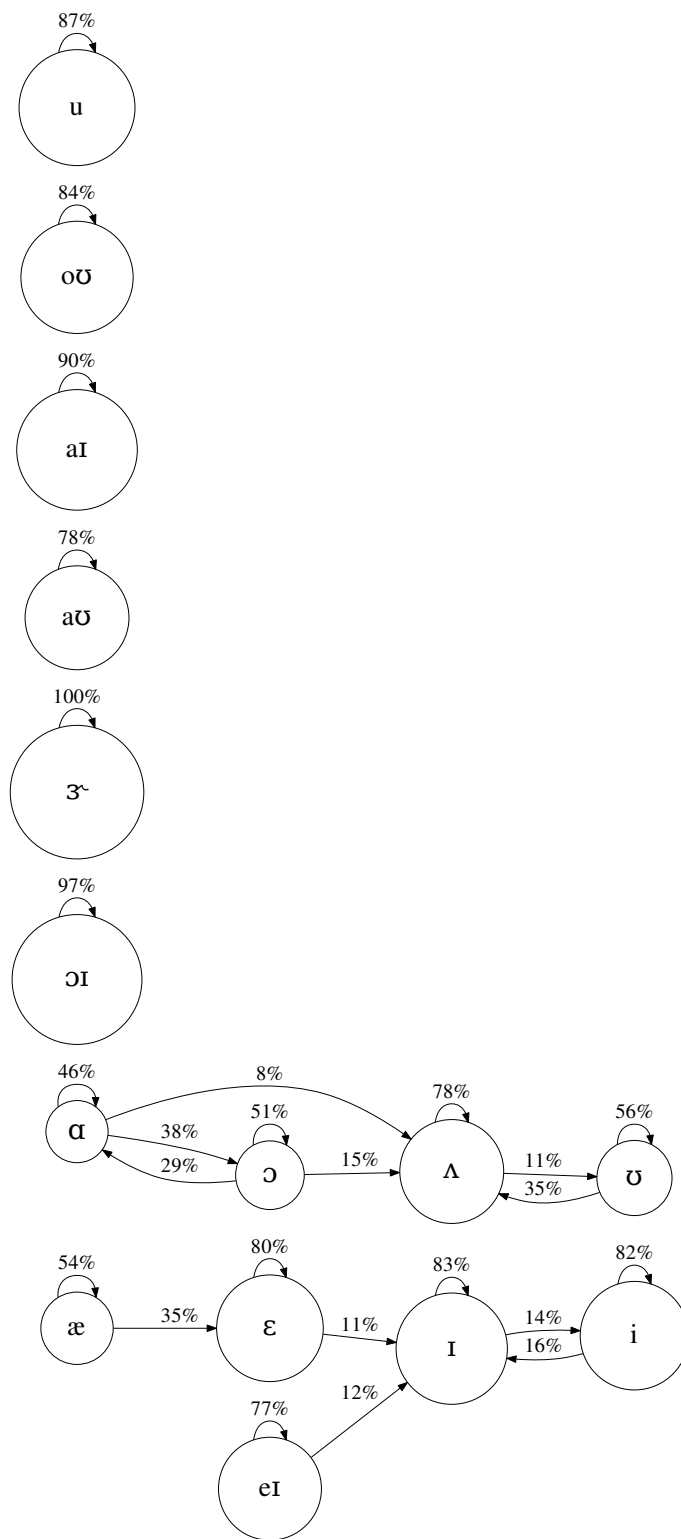
Figure E.34: Evan (RC) VI
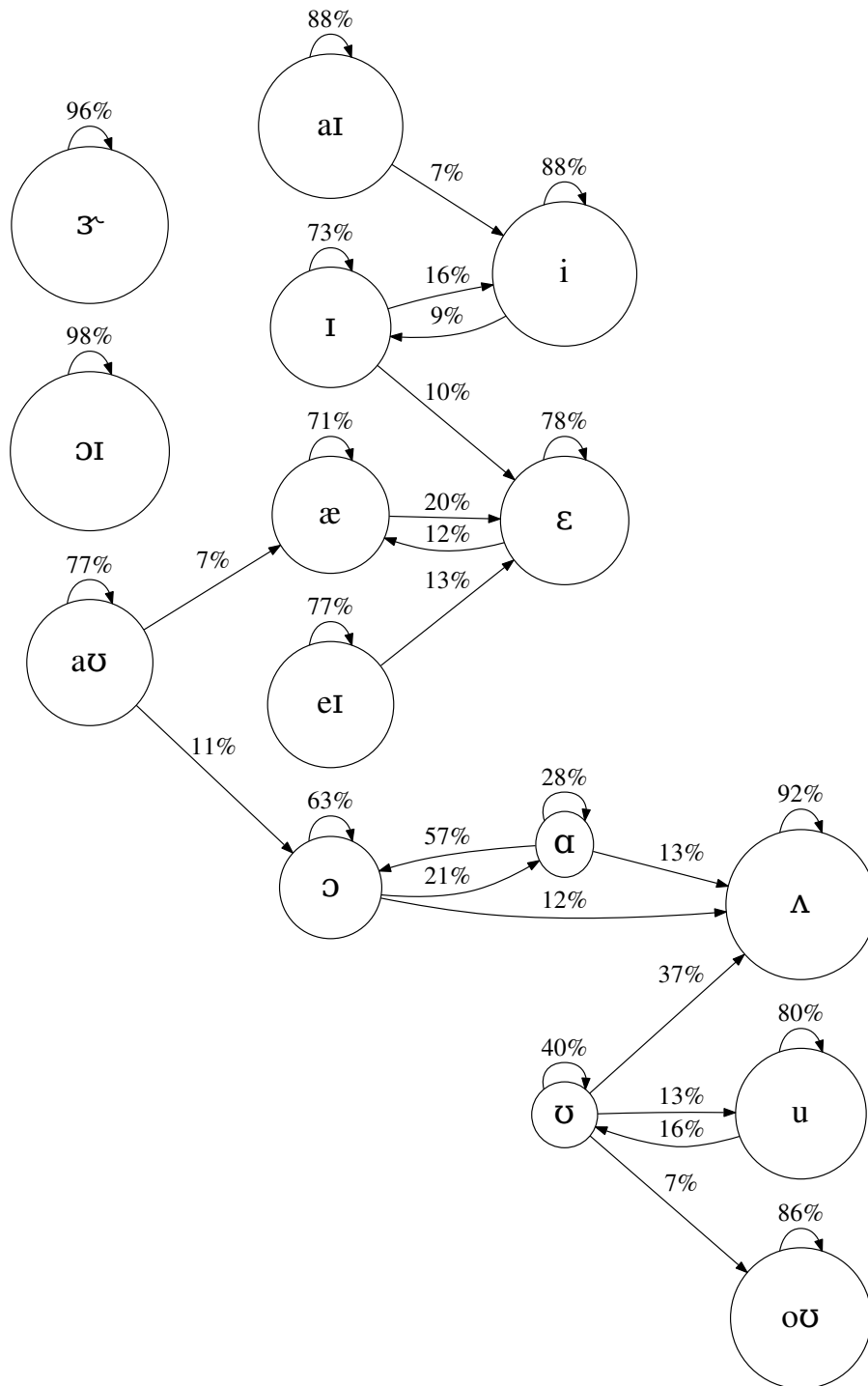
Figure E.35: Joanna (RC) VI
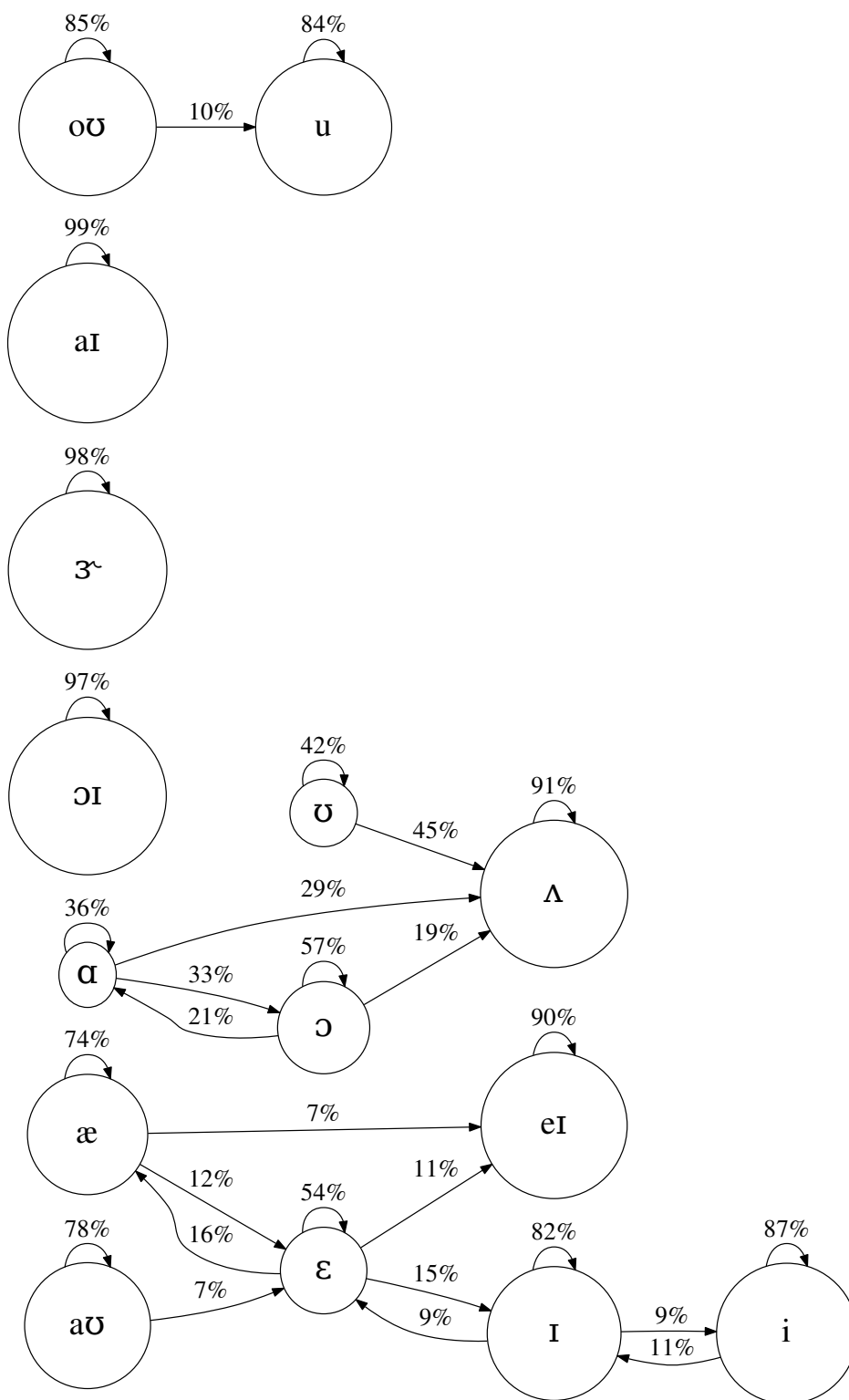
Figure E.36: Miguel (RC) VI

129
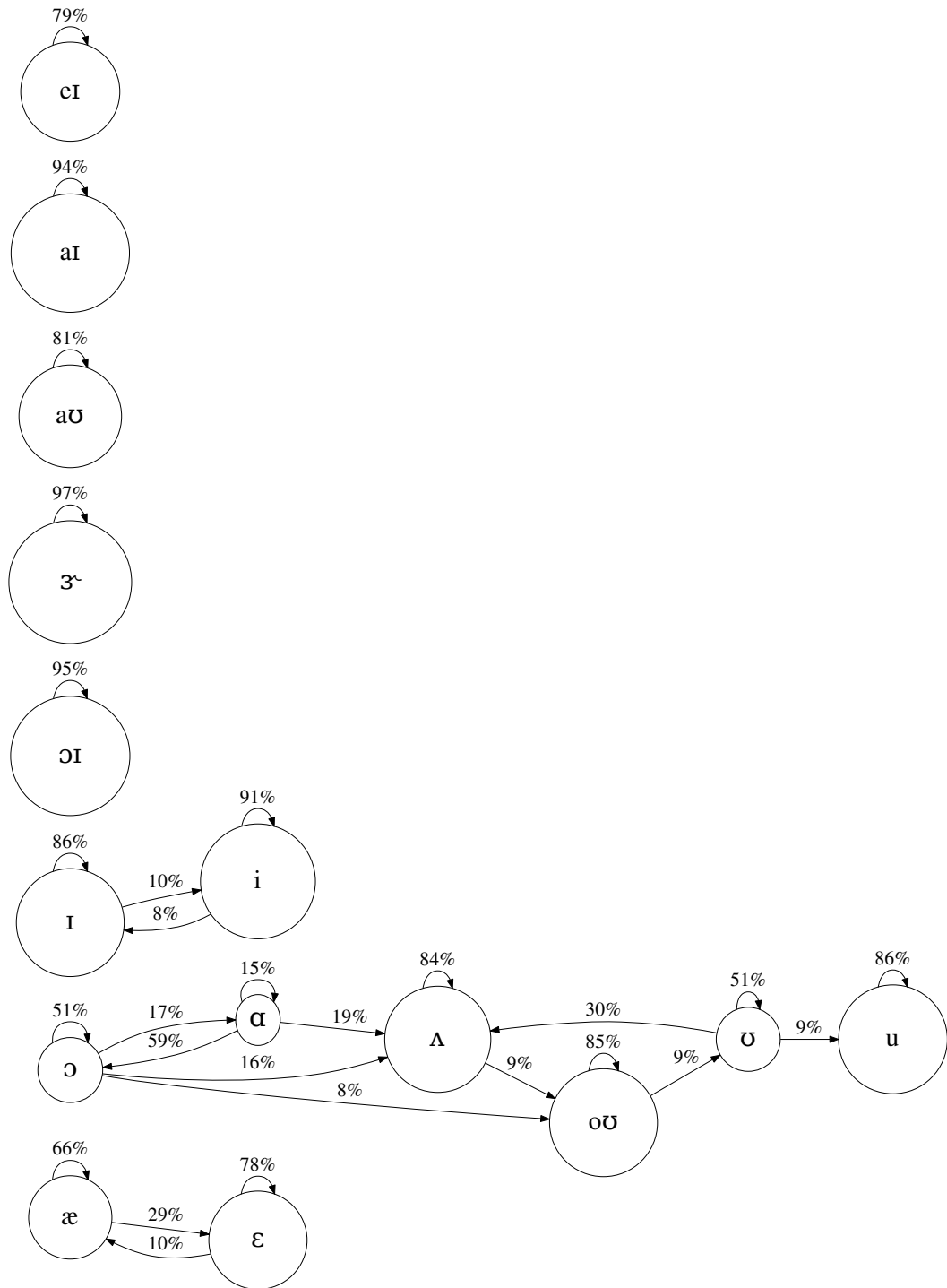
Figure E.37: Alina (RD) VI
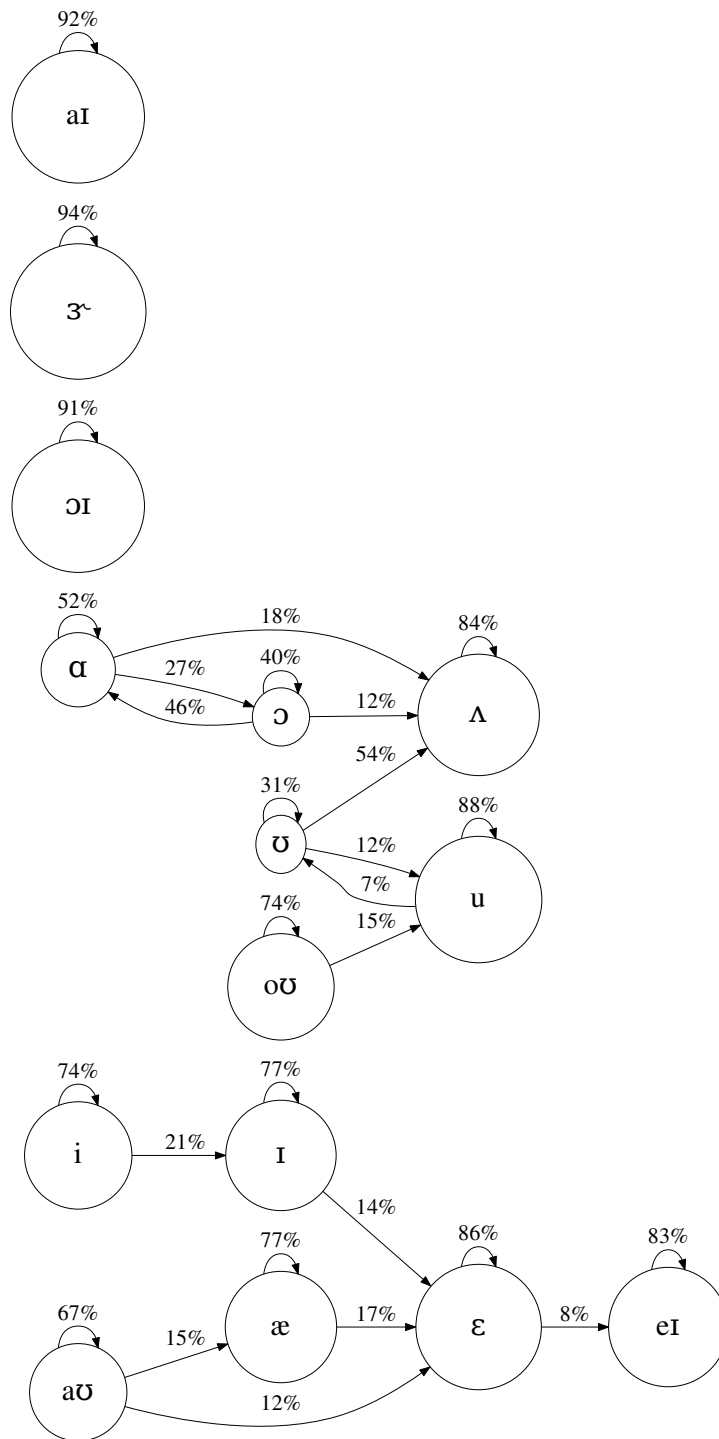
Figure E.38: Angela (RD) VI

131
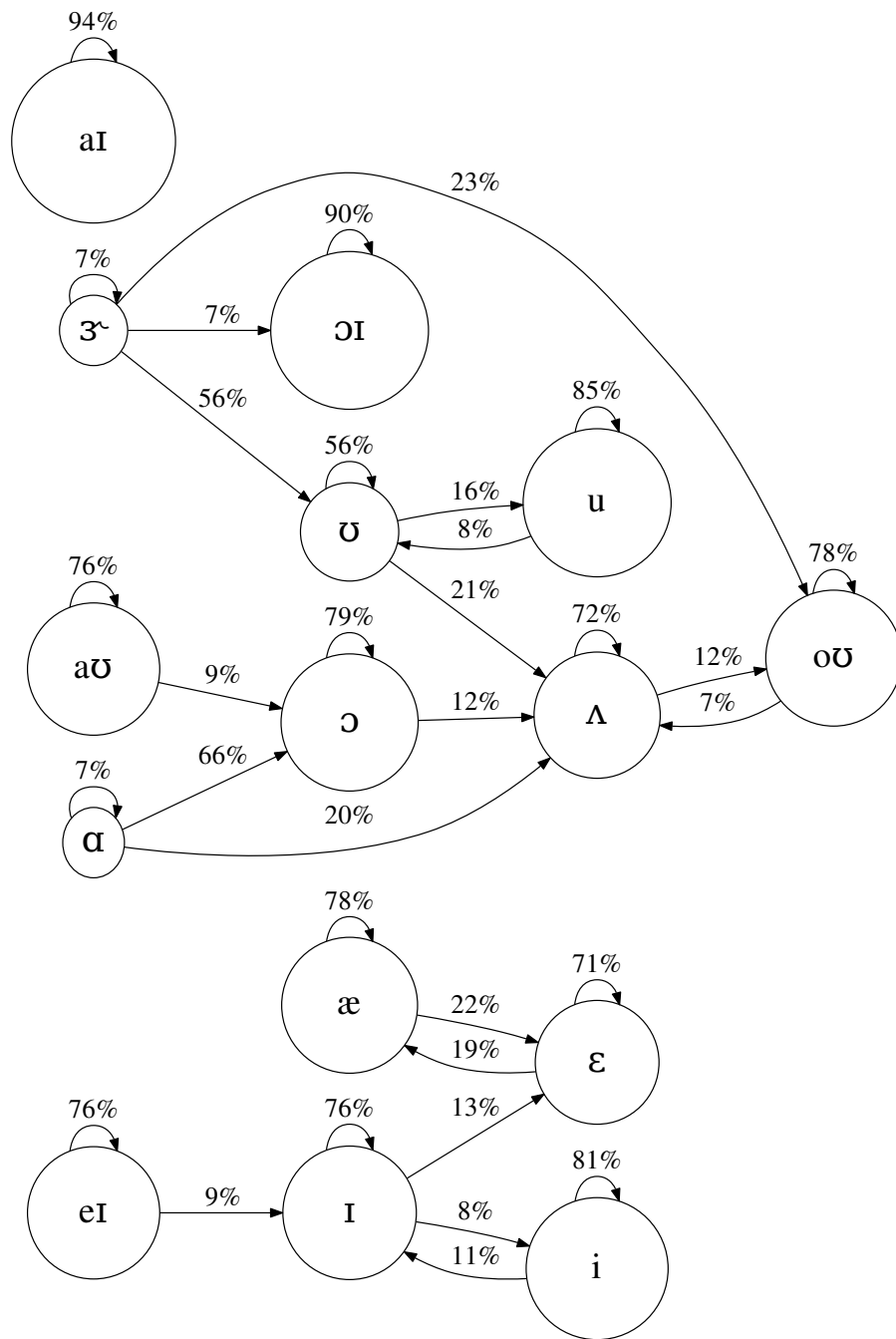
Figure E.39: Edward (RD) VI. Note that Edward could not articulate rhotic phones /r/ or /ɝ/ in his GFTA-2 articulation test. Readers of this graph should ignore any confusion patterns related to these phones.

Figure E.40: Latisha (RD) VI

133
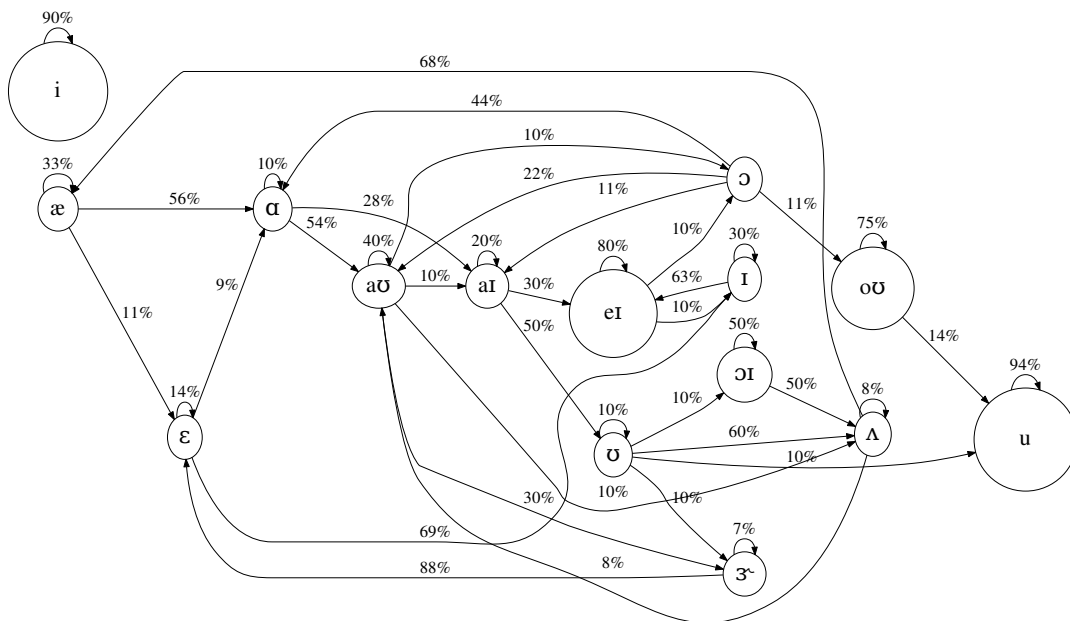
Figure E.41: Laura (RD) VI
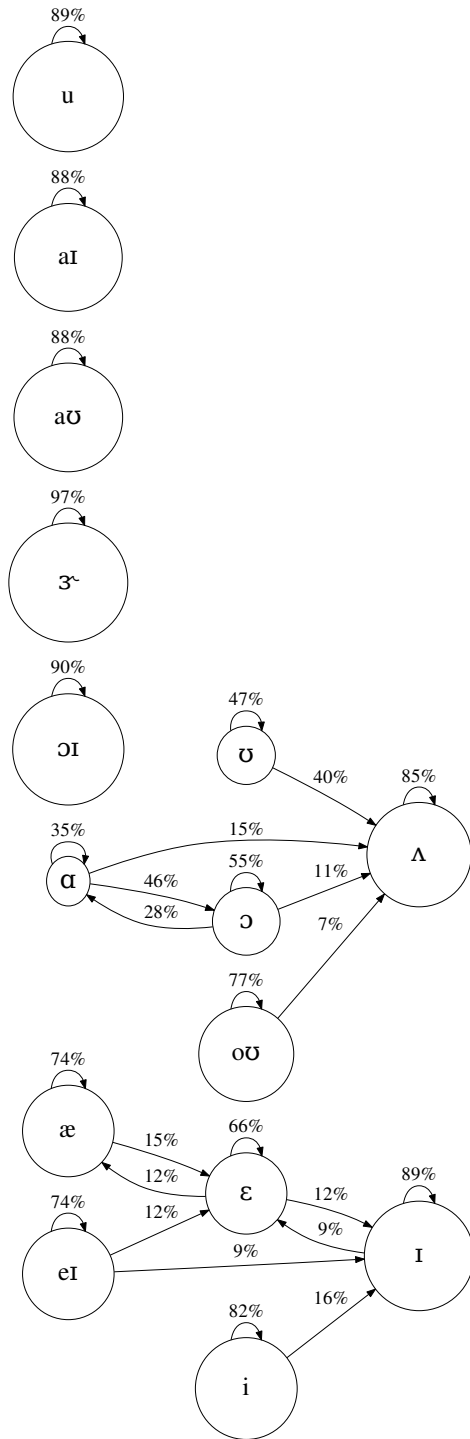
Figure E.42: Norene (RD) VI

135
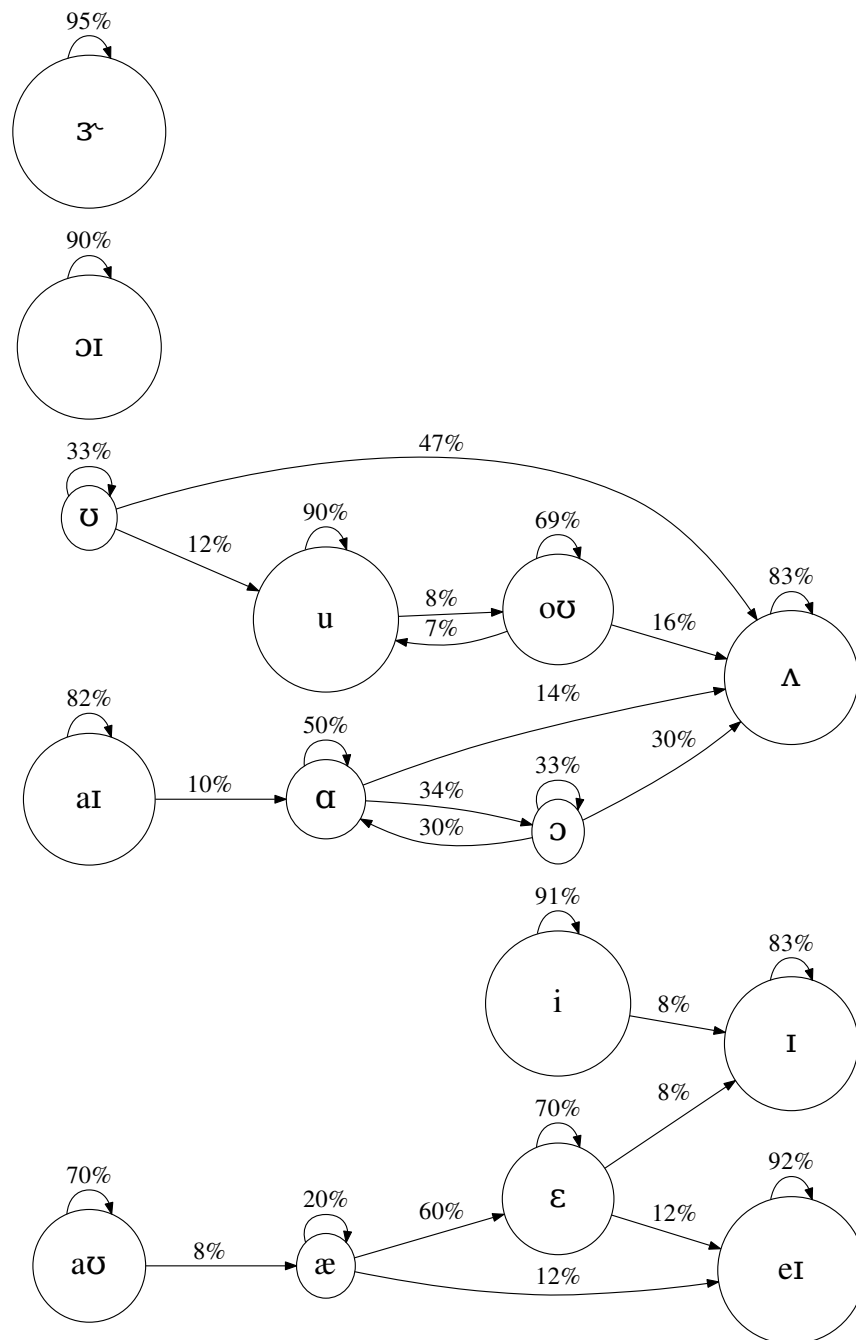
Figure E.43: Shauna (RD) VI
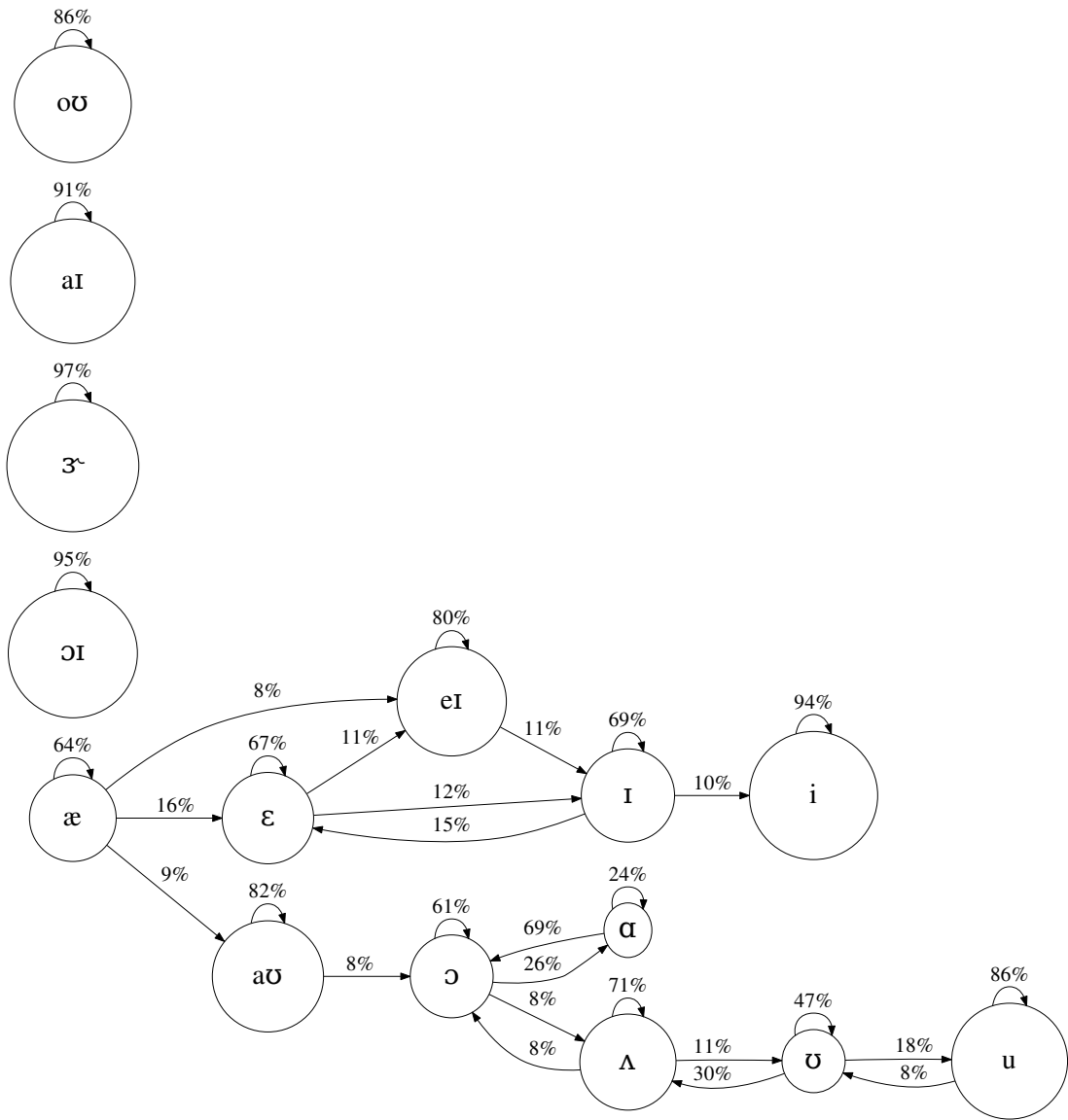
Figure E.44: Teddy (RD) VI

137
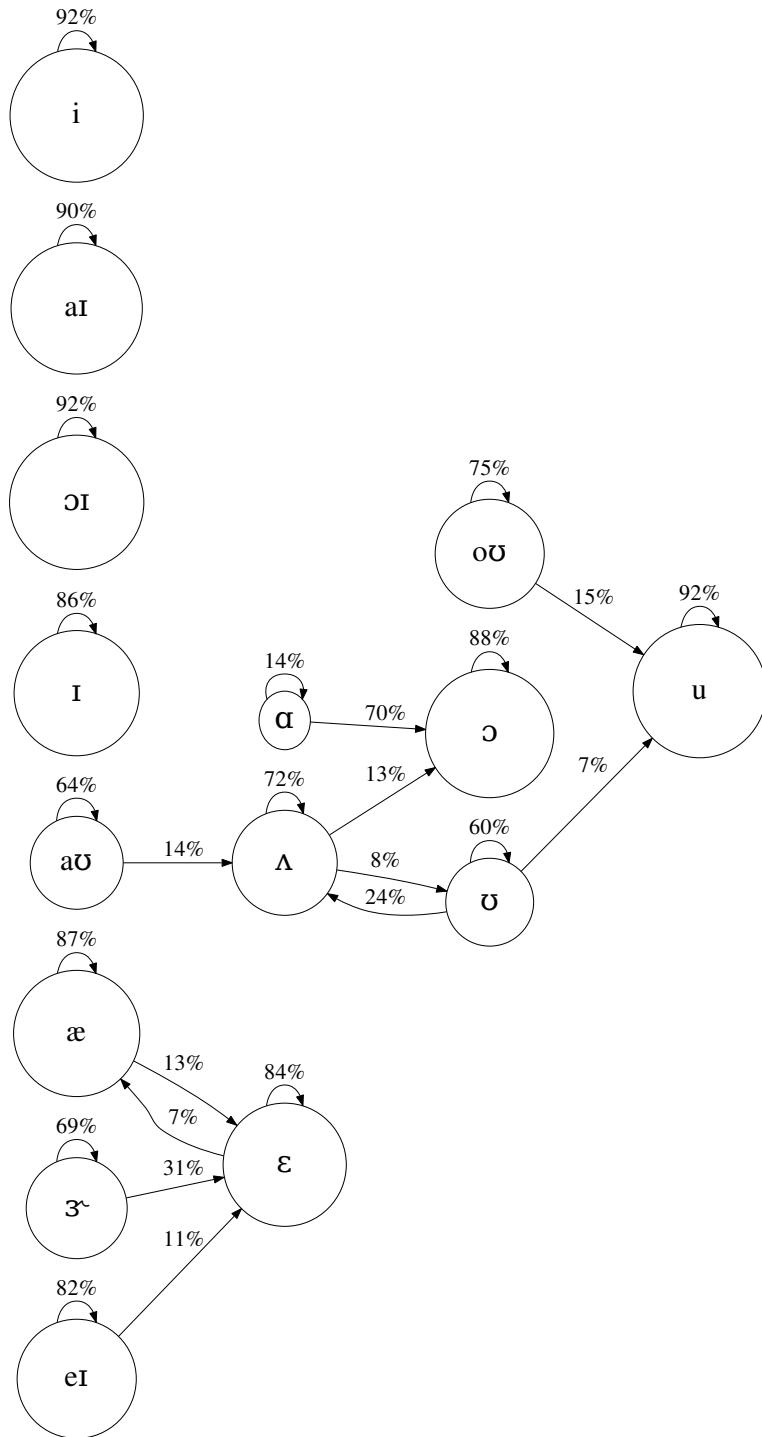
Figure E.45: Tony (RD) VI
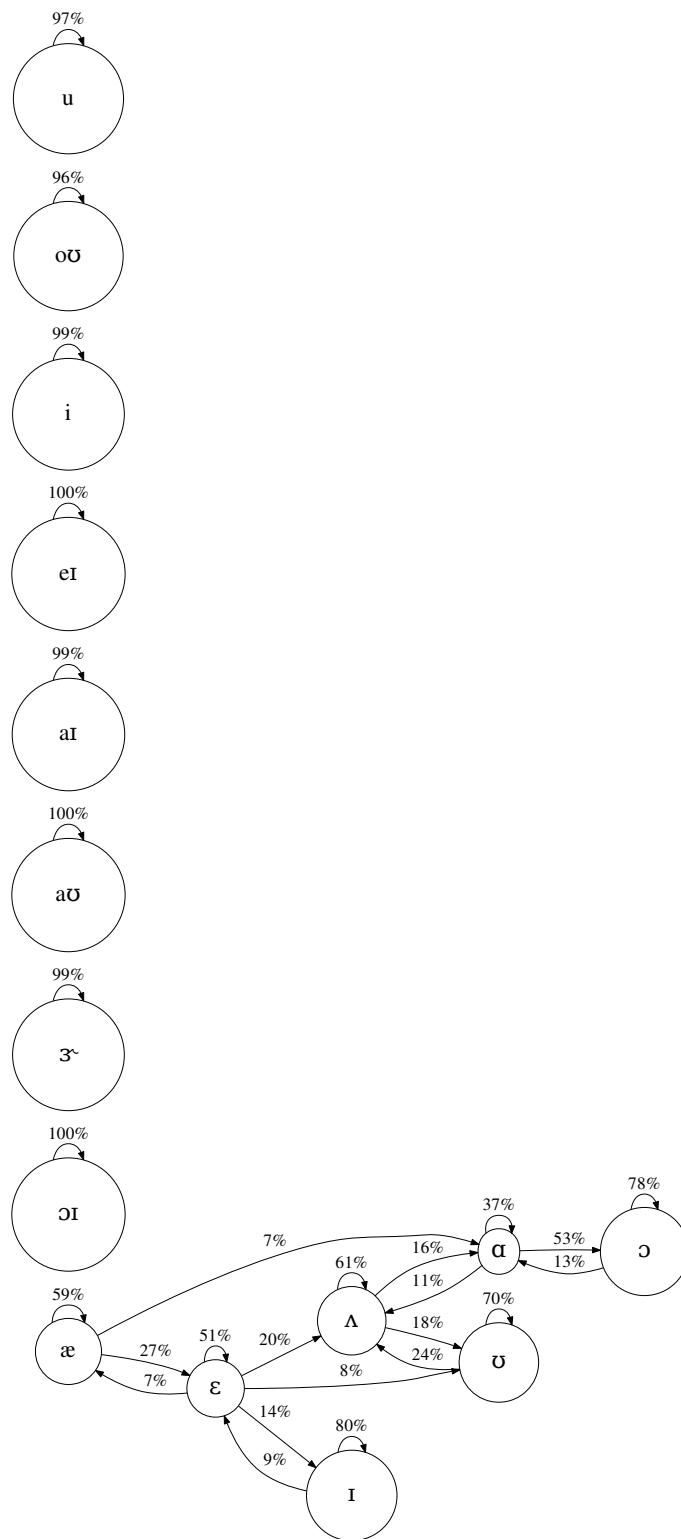
Figure E.46: Anton (RC) VF

139
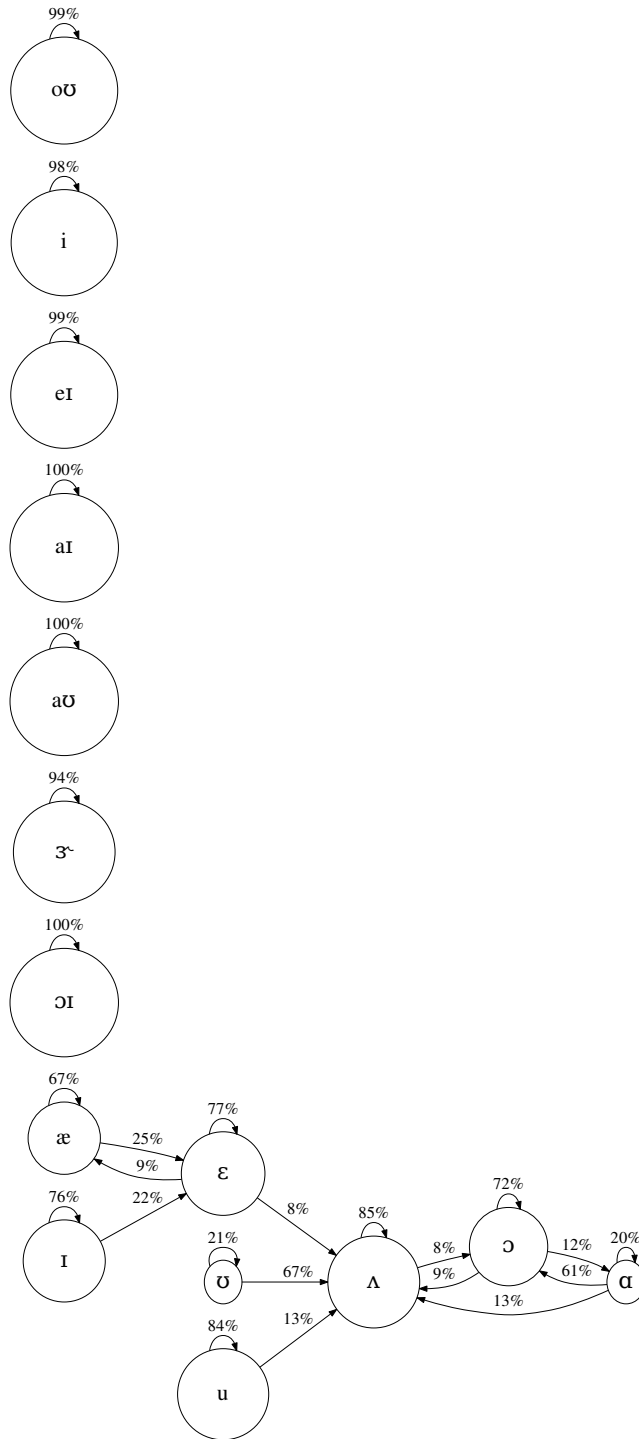
Figure E.47: Bob (RC) VF

140
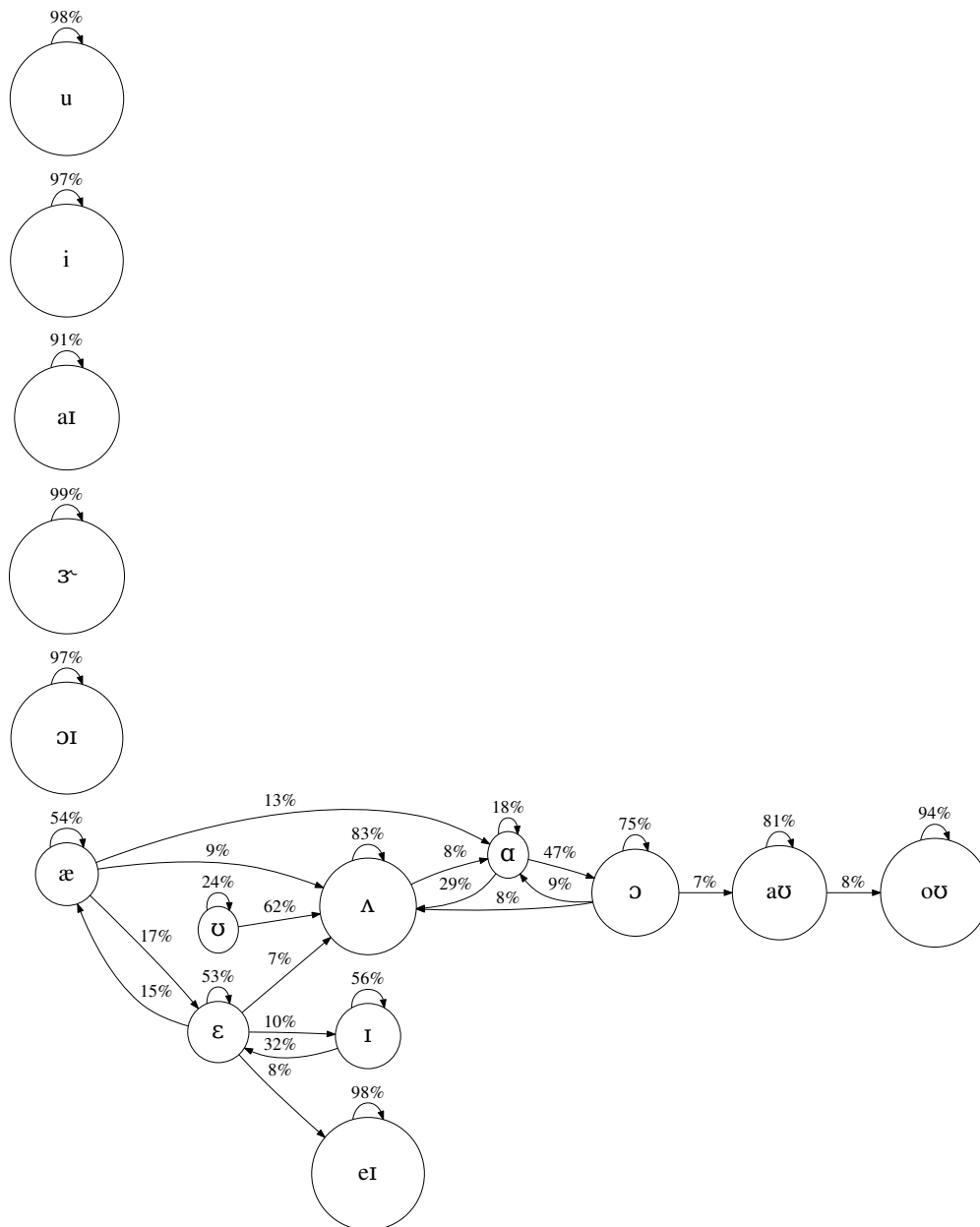
Figure E.48: Carly (RC) VF
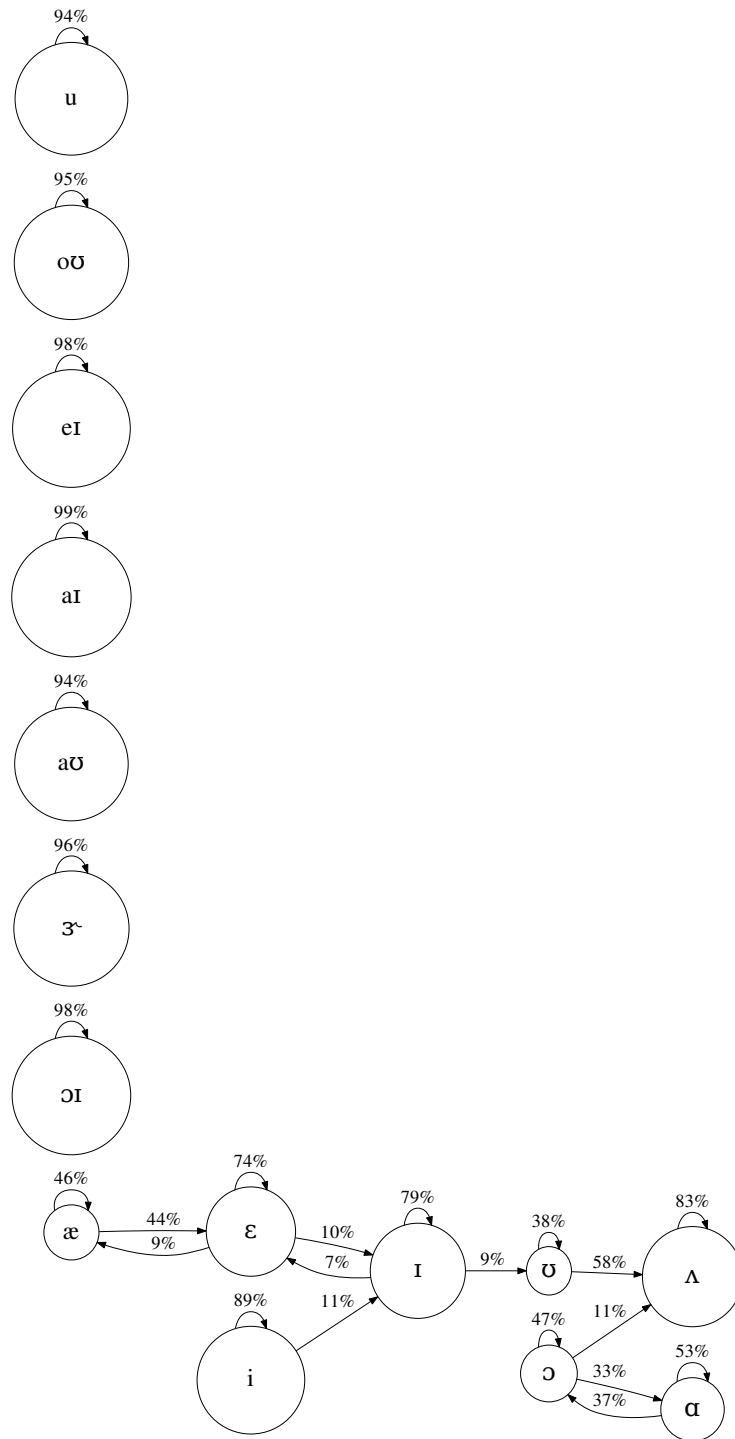
Figure E.49: Evan (RC) VF
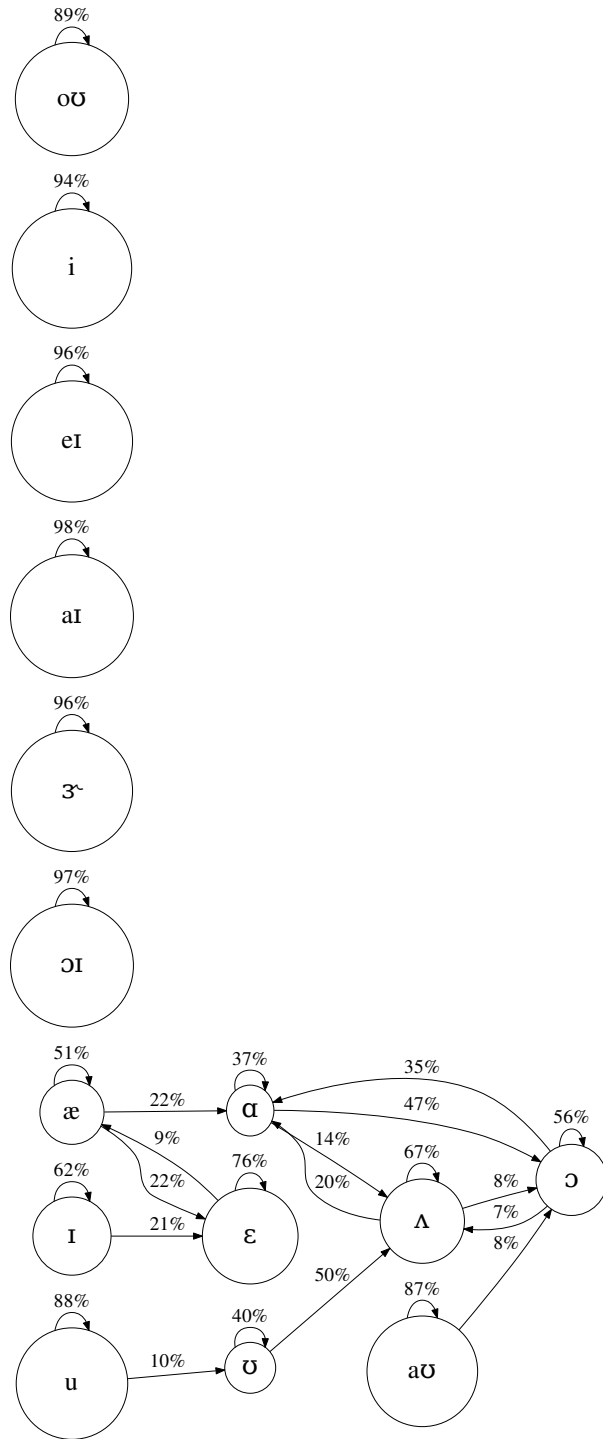
Figure E.50: Joanna (RC) VF

143
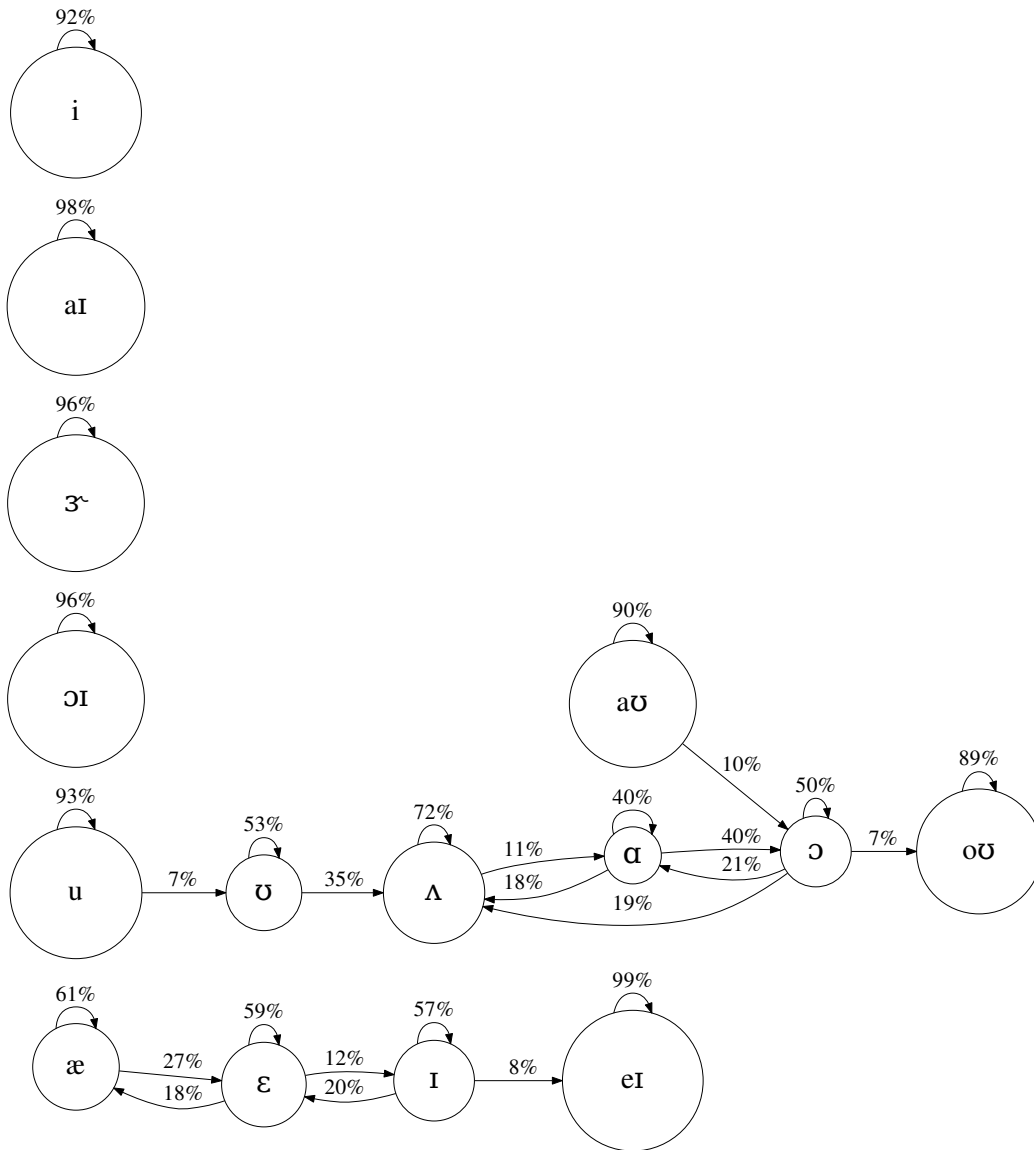
Figure E.51: Miguel (RC) VF

144
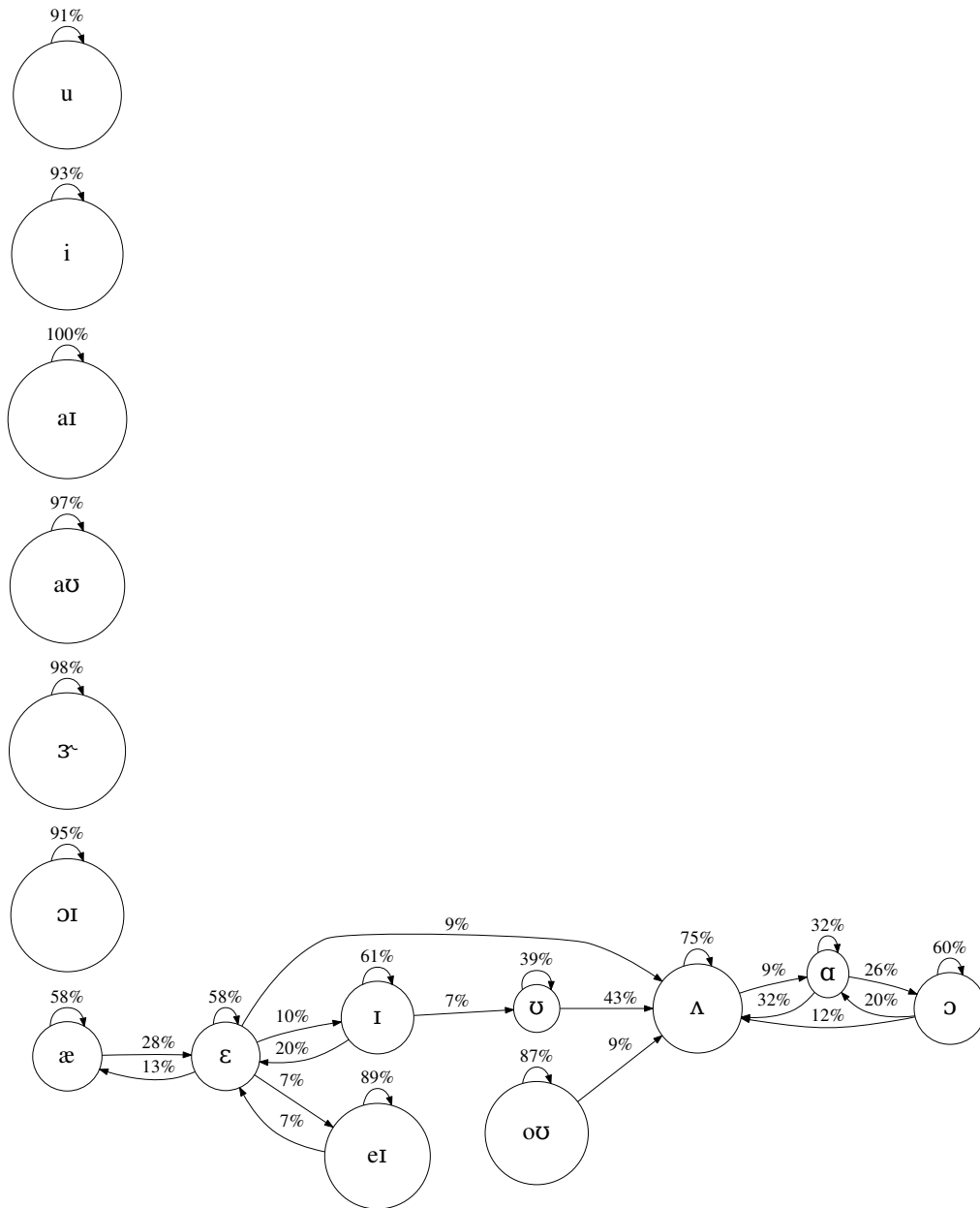
Figure E.52: Alina (RD) VF

145
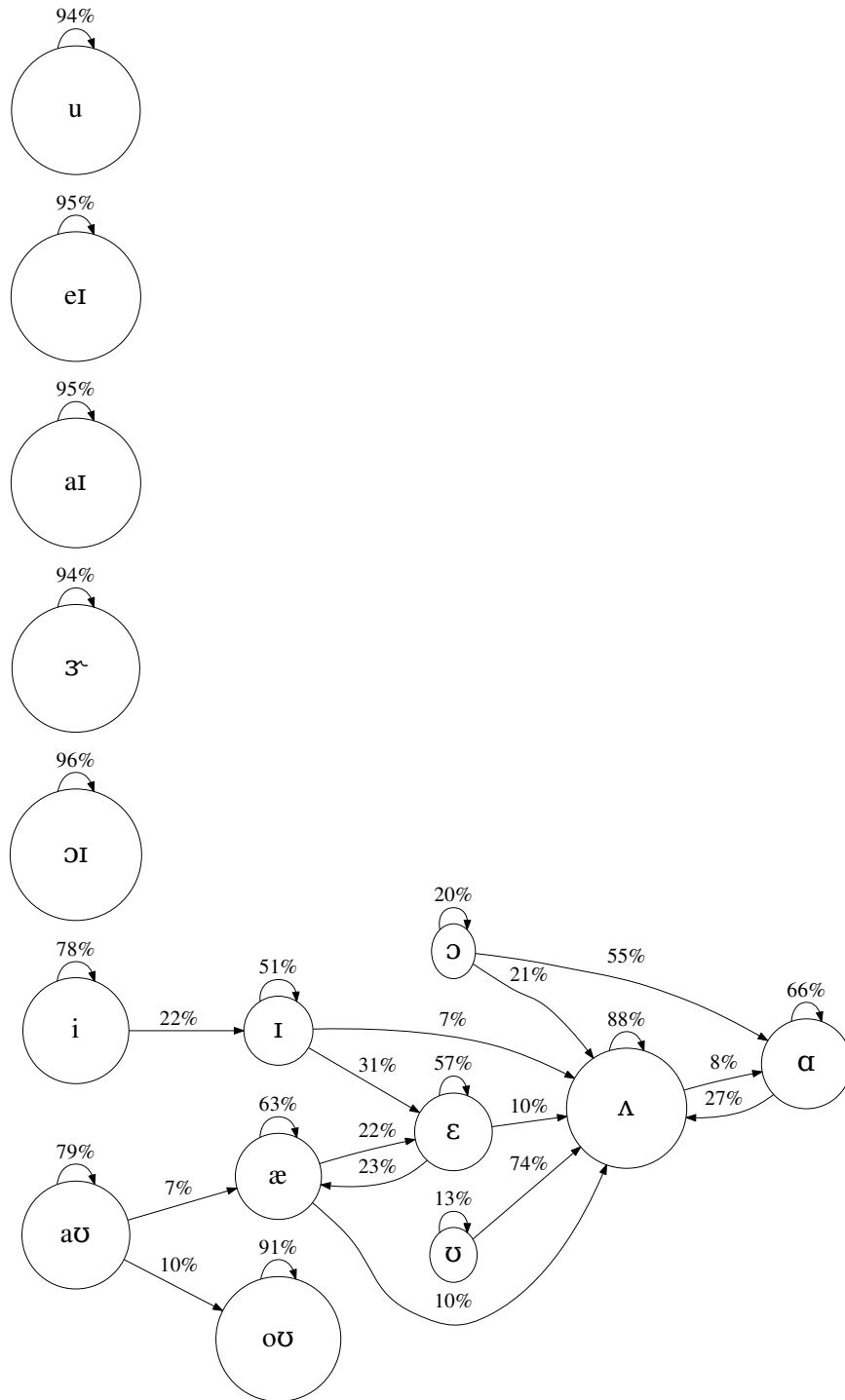
Figure E.53: Angela (RD) VF

146
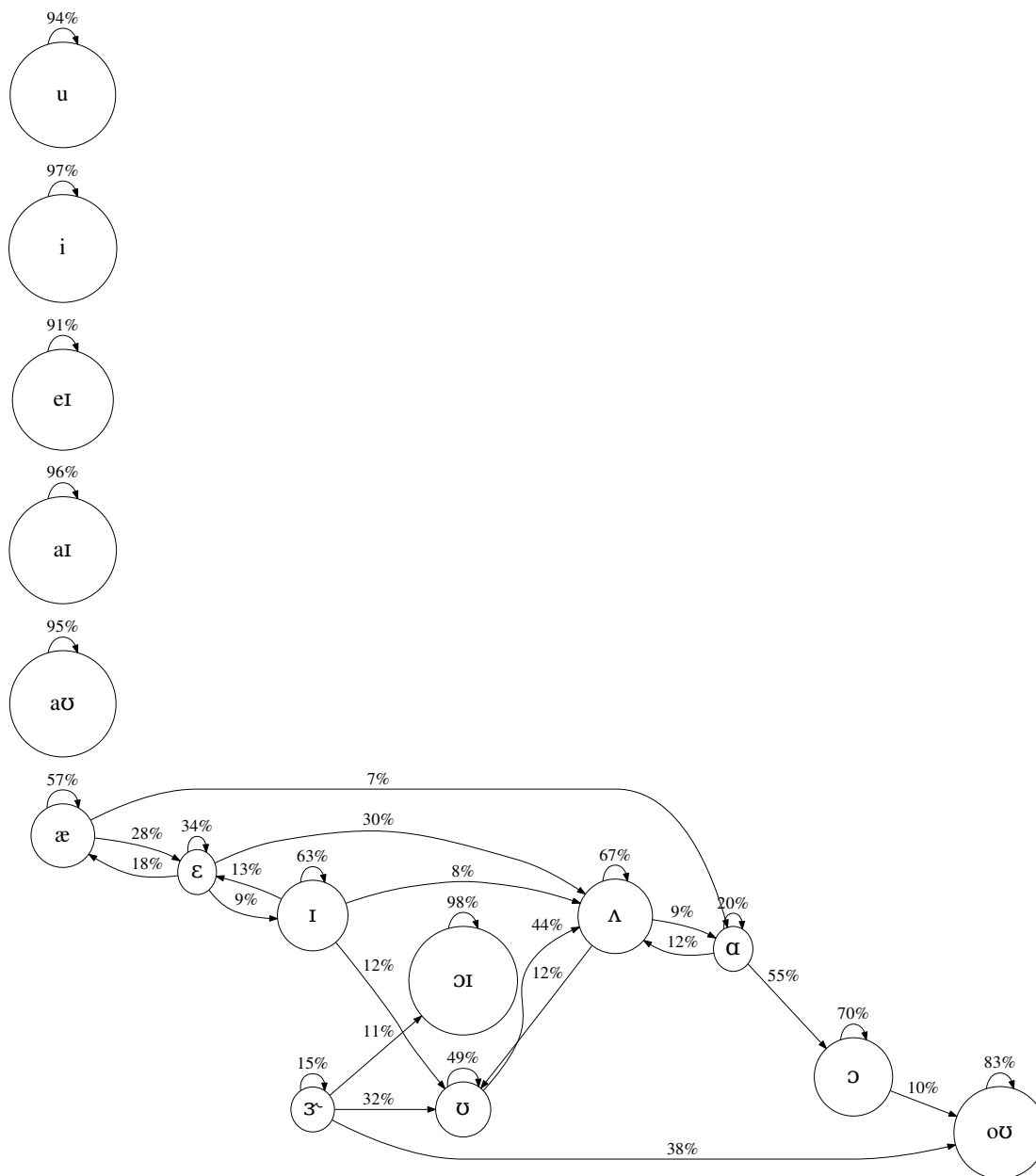
Figure E.54: Edward (RD) VF. Note that Edward could not articulate rhotic phones /r/ or /ɜ/ in his GFTA-2 articulation test. Readers of this graph should ignore any confusion patterns related to these phones.
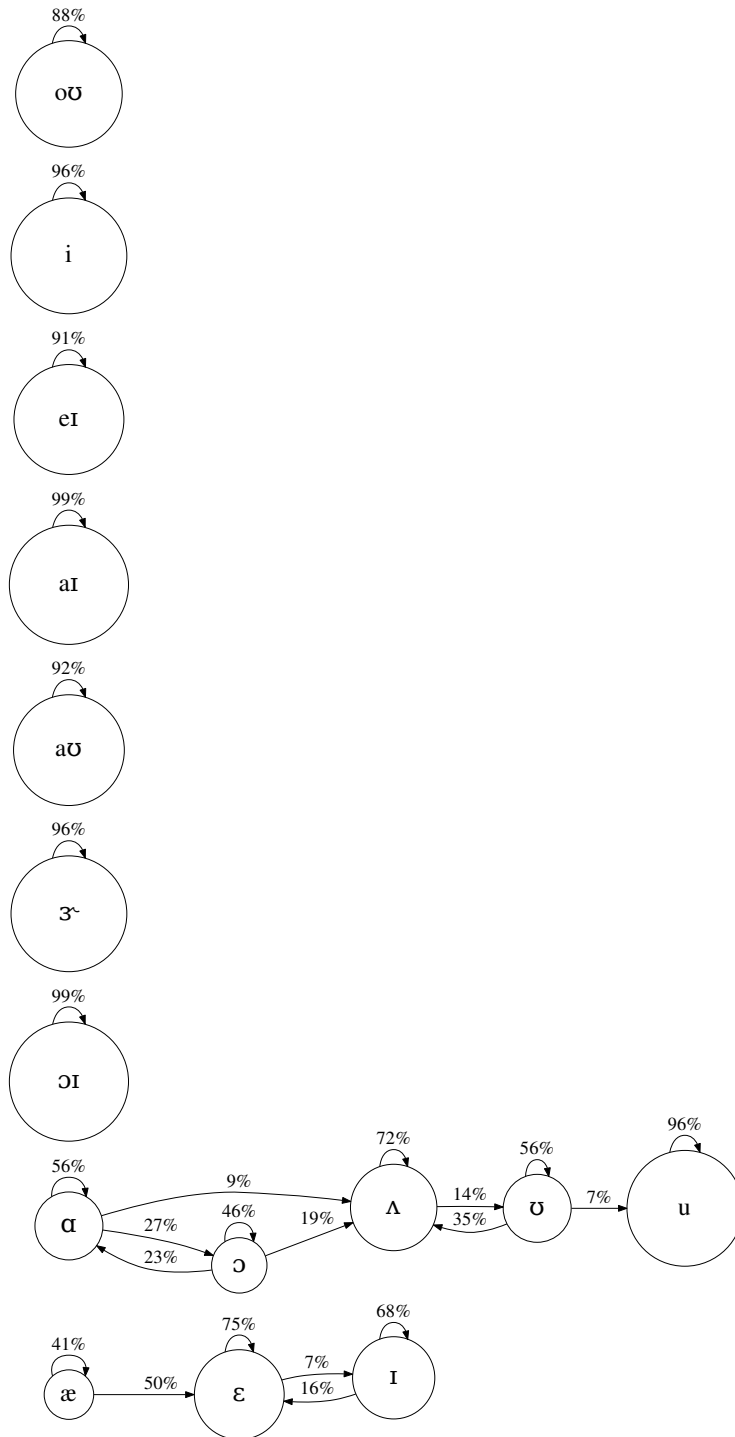
Figure E.55: Latisha (RD) VF
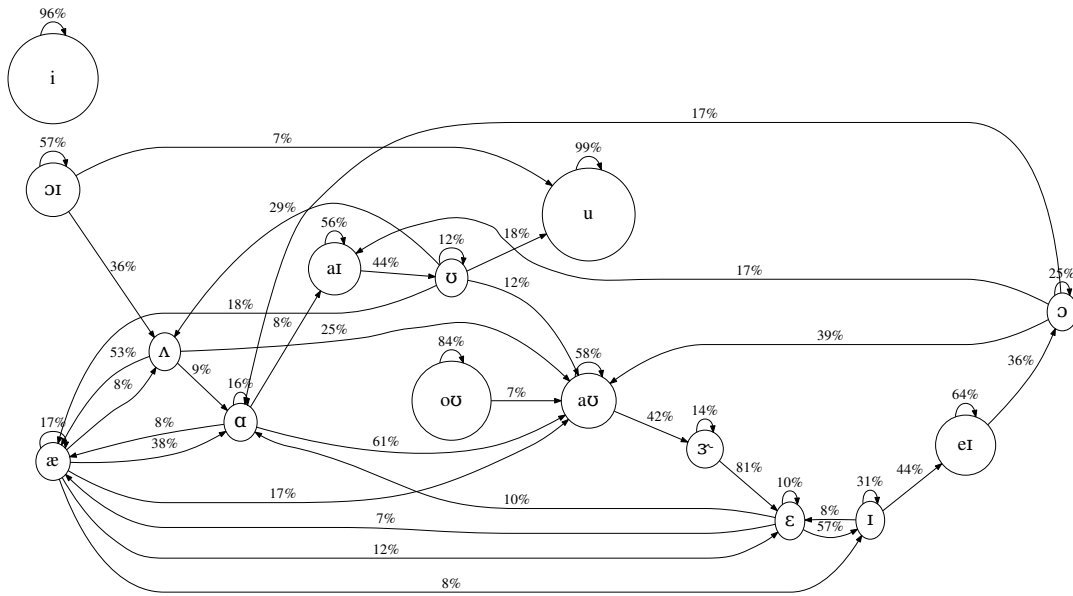
Figure E.56: Laura (RD) VF
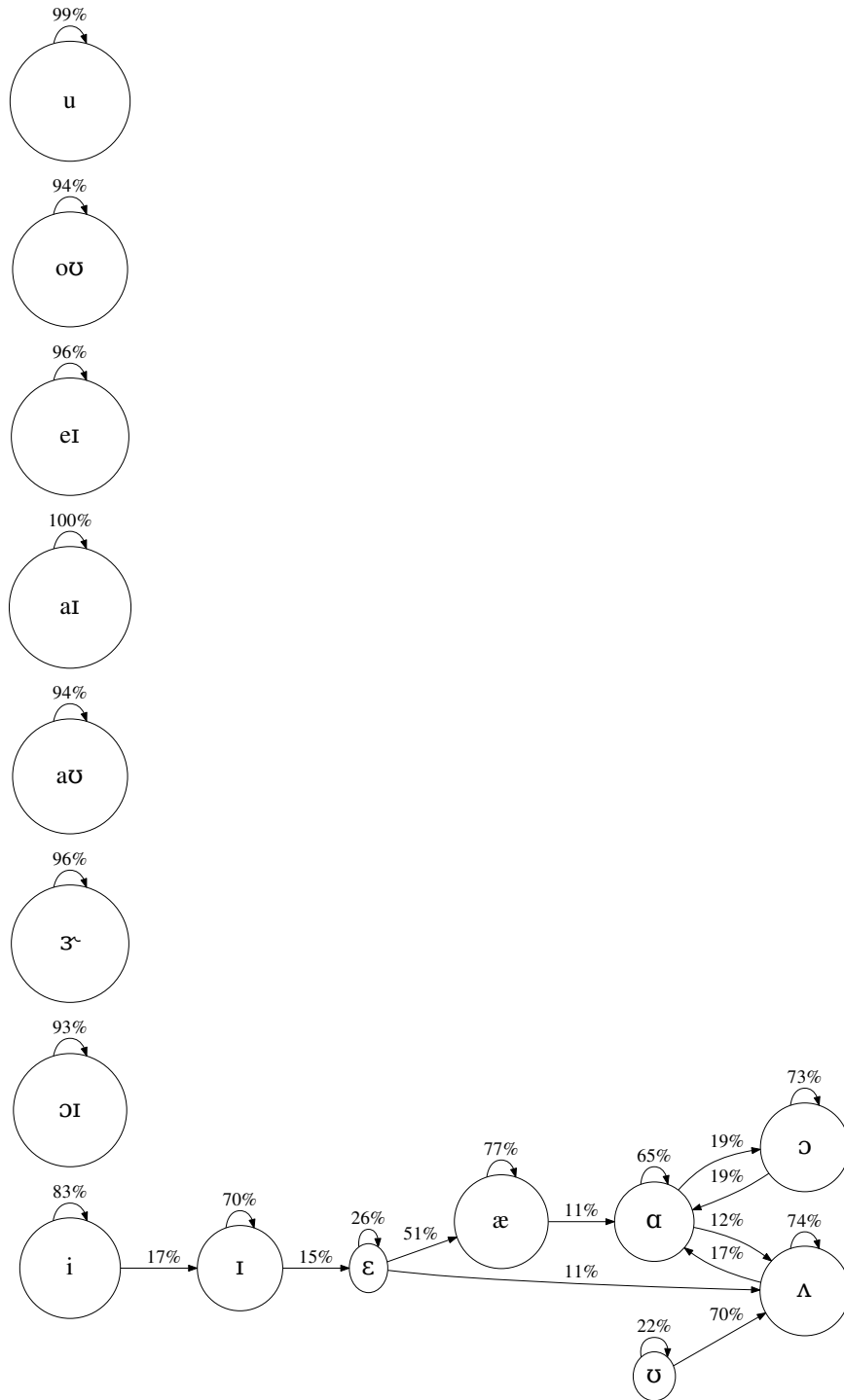
Figure E.57: Norene (RD) VF

150
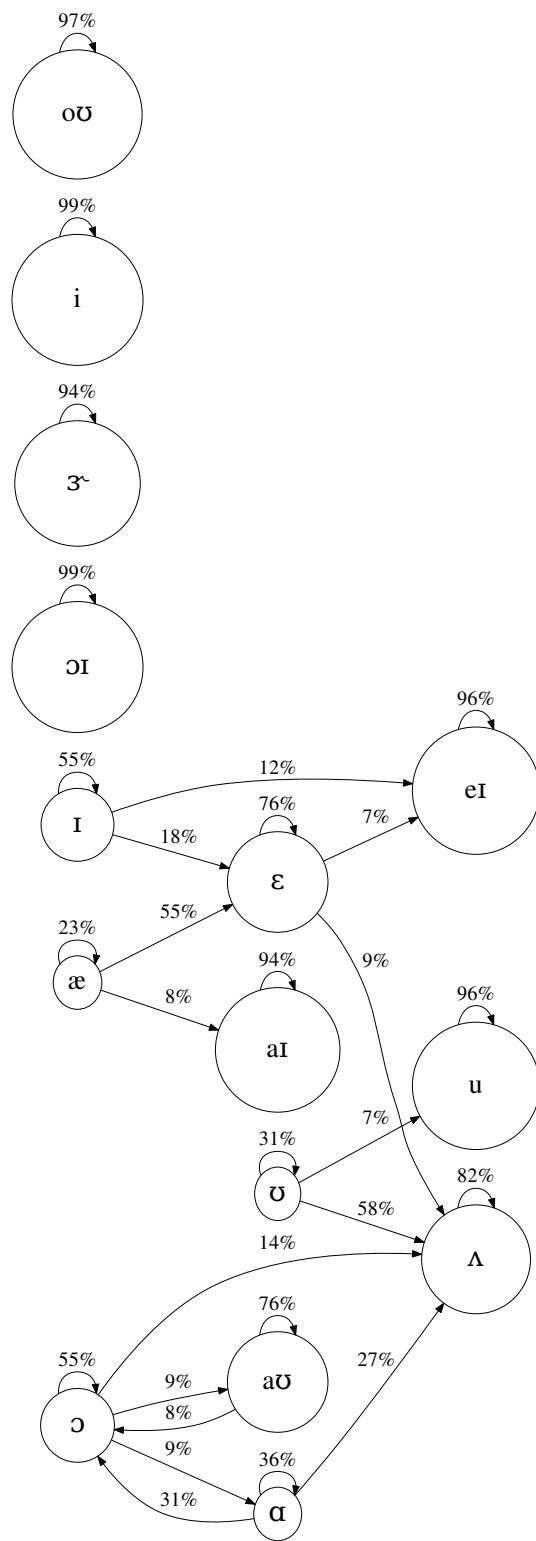
Figure E.58: Shauna (RD) VF

151
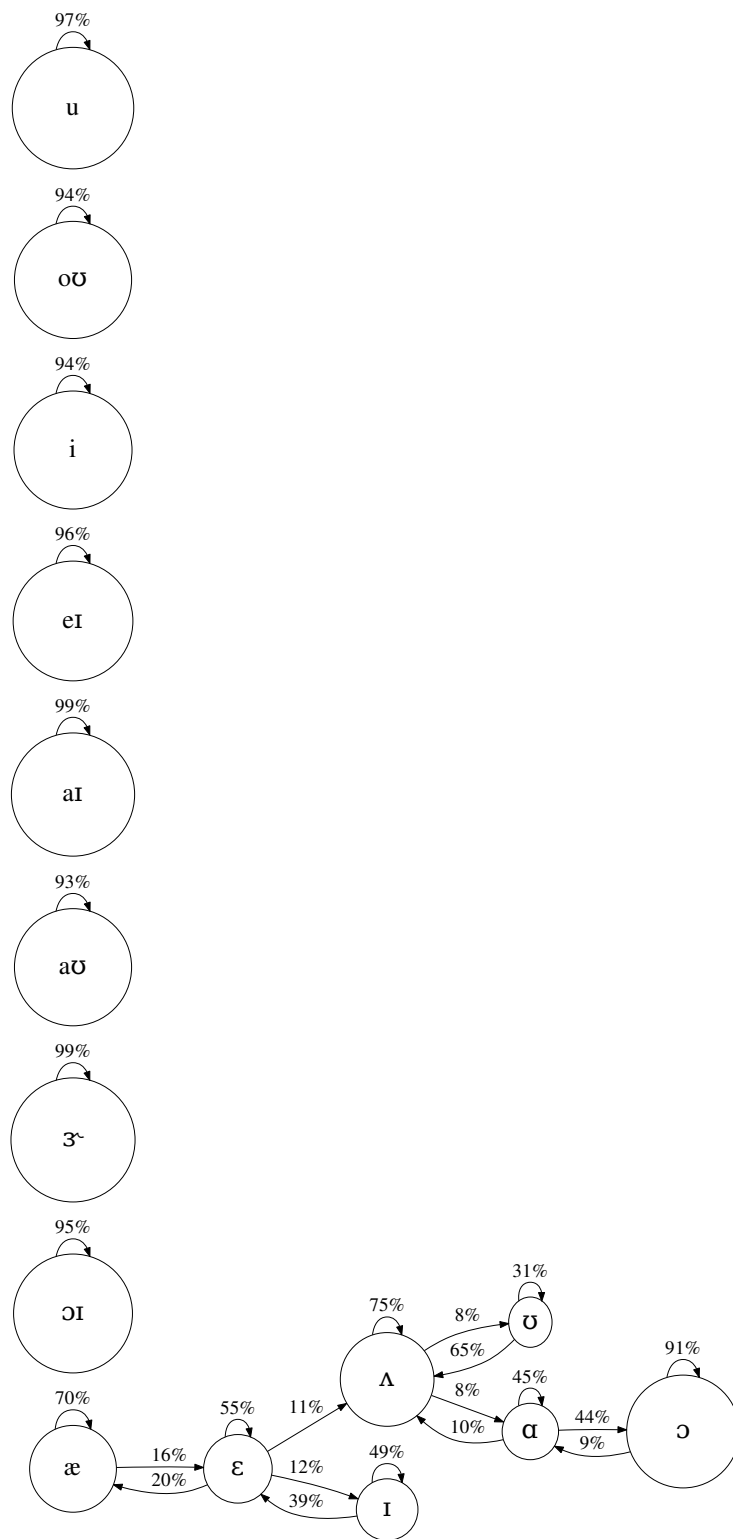
Figure E.59: Teddy (RD) VF

152
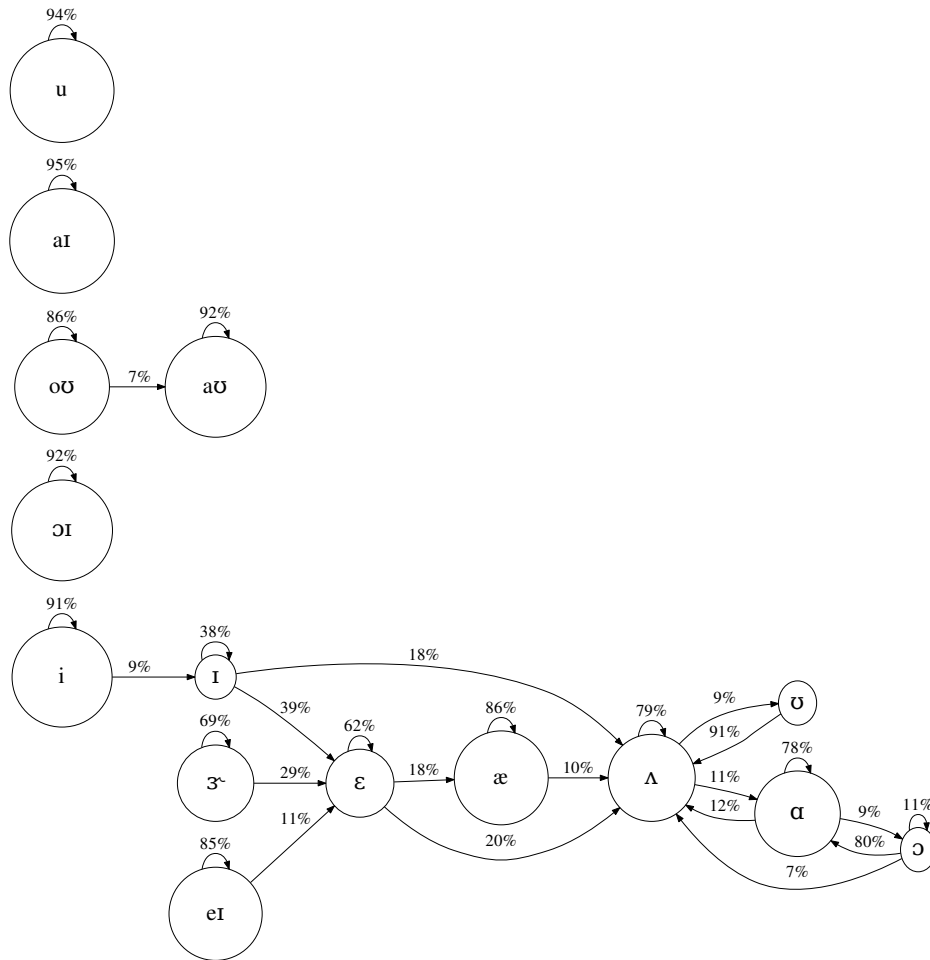
Figure E.60: Tony (RD) VF

153

Appendix F

REORDERING WITH PERMUTATION
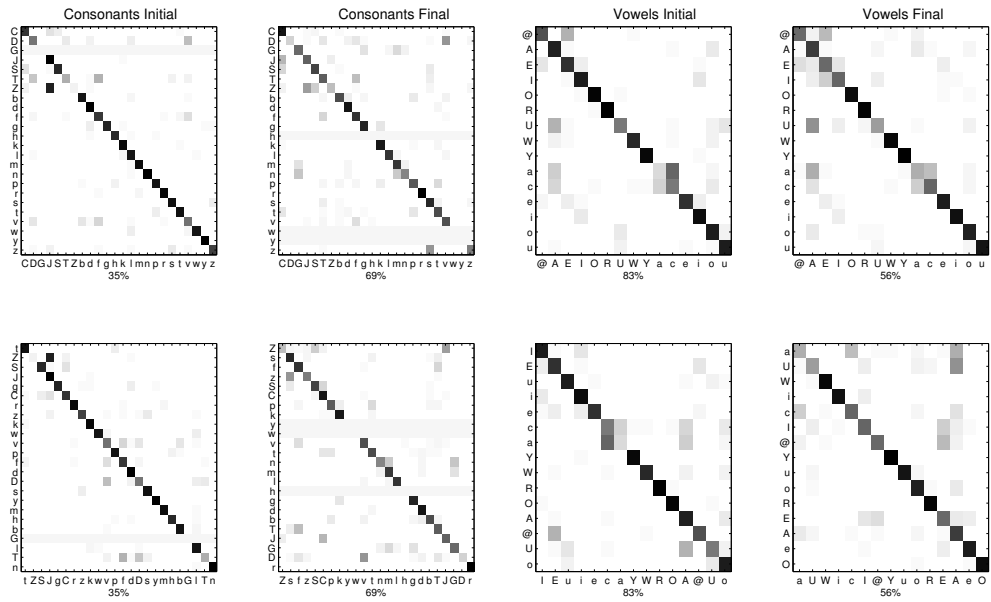
(a) RD-Alina
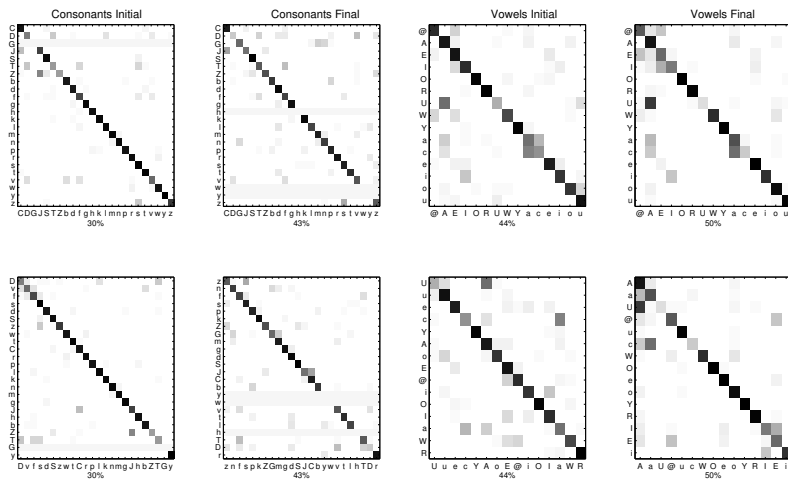
(b) RD-Angela

Figure F.1: Diagonalization of confusion matrices for RD subjects

Consonants Initial · Consonants Final · Vowels Initial · Vowels Final

(c) RD-Edward

Consonants Initial · Consonants Final · Vowels Initial · Vowels Final
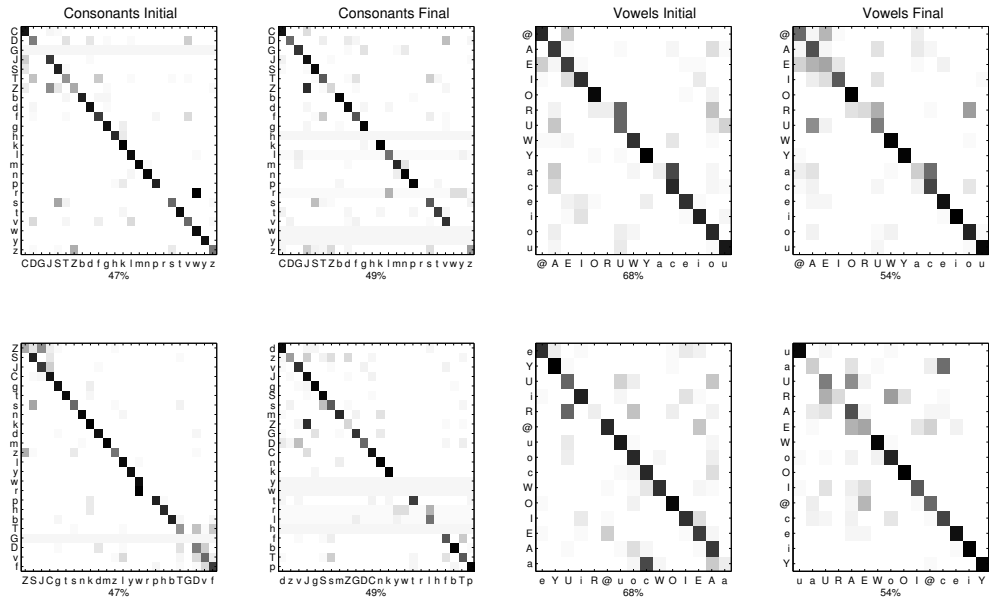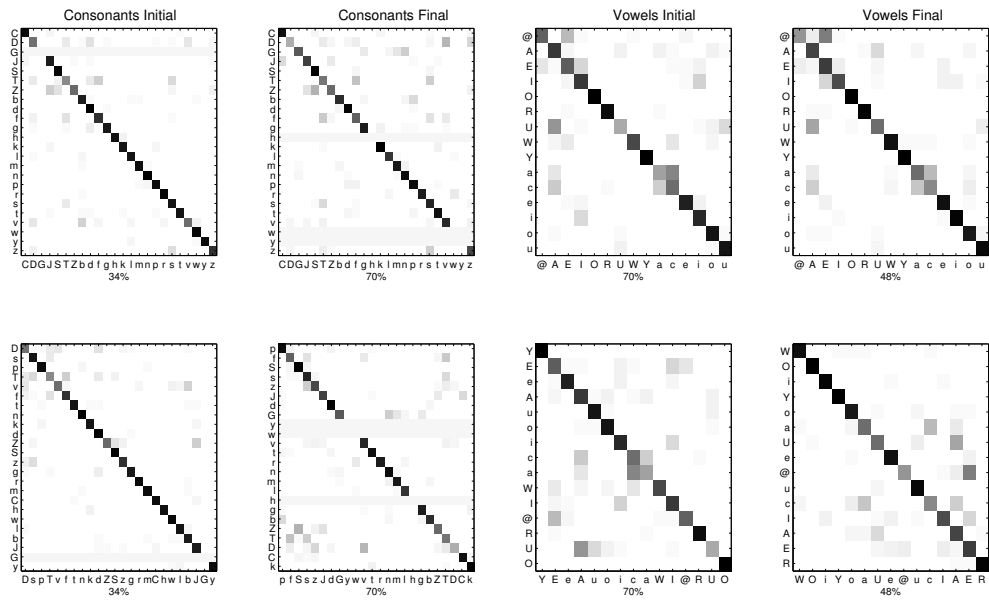
(d) RD-Latisha

Figure F.1: cont.

155
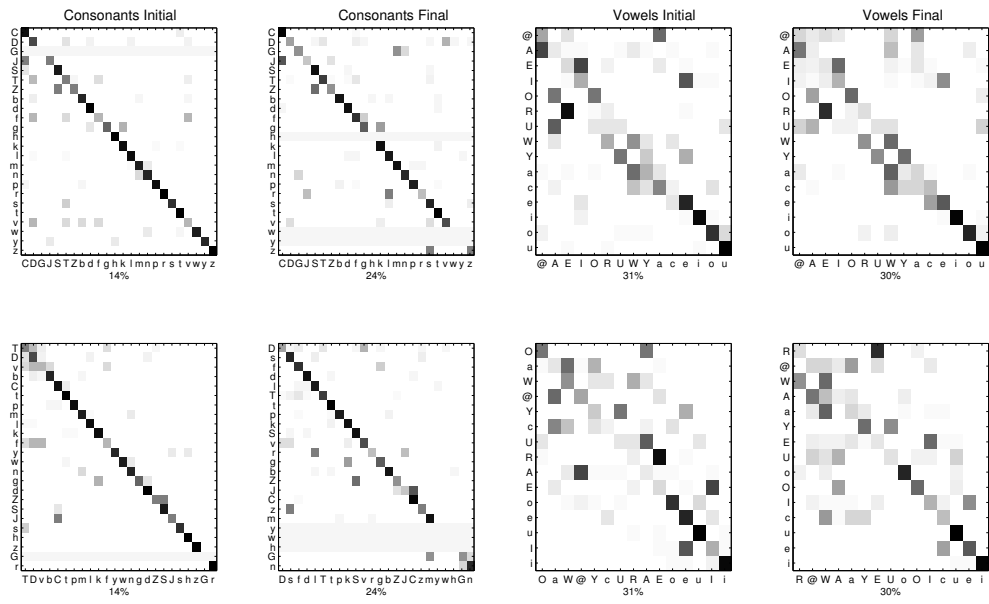
(e) RD-Laura
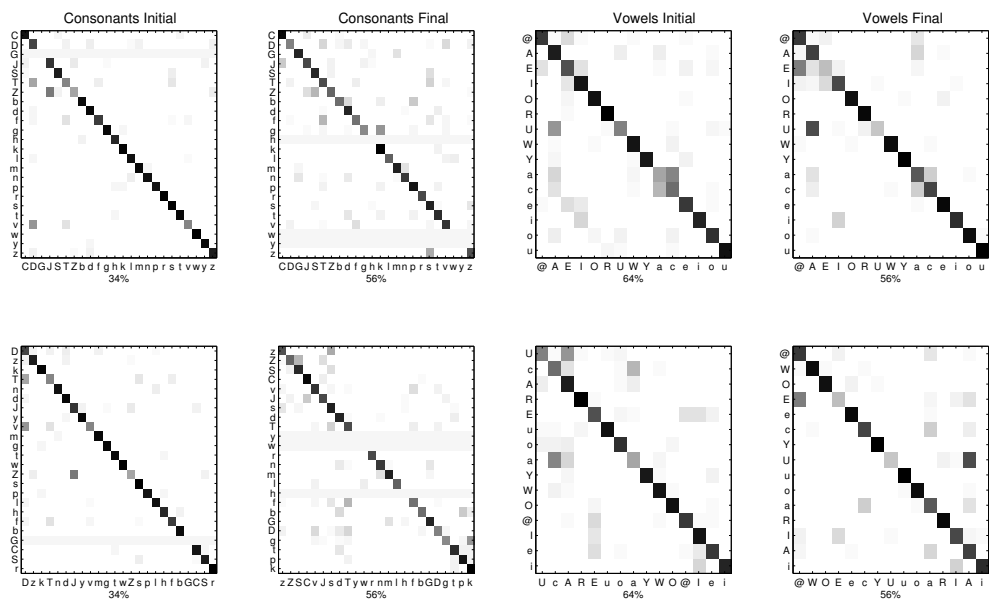
(f) RD-Norene

Figure F.1: cont.

Consonants Initial 38% Consonants Final 50% Vowels Initial 59% Vowels Final 62%

Consonants Initial 38% Consonants Final 50% Vowels Initial 59% Vowels Final 62%

(g) RD-Shauna

Consonants Initial 27% Consonants Final 43% Vowels Initial 50% Vowels Final 40%

Consonants Initial 27% Consonants Final 43% Vowels Initial 50% Vowels Final 40%

(h) RD-Teddy

Figure F.1: cont.

157

Consonants Initial   Consonants Final   Vowels Initial   Vowels Final

(i) RD-Tony

Figure F.1: cont.

Consonants Initial   Consonants Final   Vowels Initial   Vowels Final

(a) RC-Anton

Consonants Initial   Consonants Final   Vowels Initial   Vowels Final

(b) RC-Bob

Figure F.2: Diagonalization of confusion matrices for RC subjects
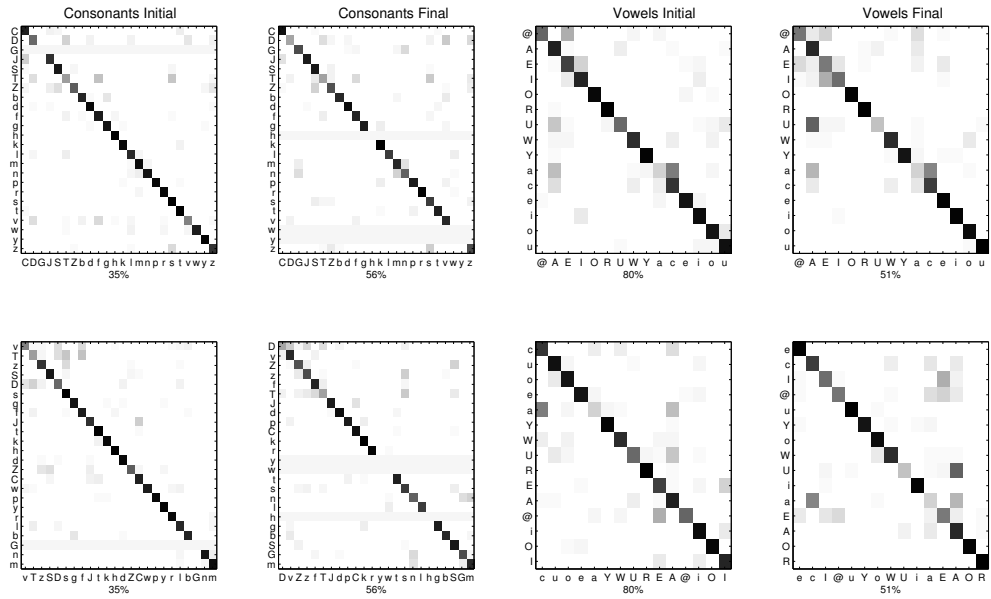
(c) RC-Carly

(d) RC-Evan

Figure F.2: cont.

Consonants Initial   Consonants Final   Vowels Initial   Vowels Final

(e) RC-Joanna

Consonants Initial   Consonants Final   Vowels Initial   Vowels Final
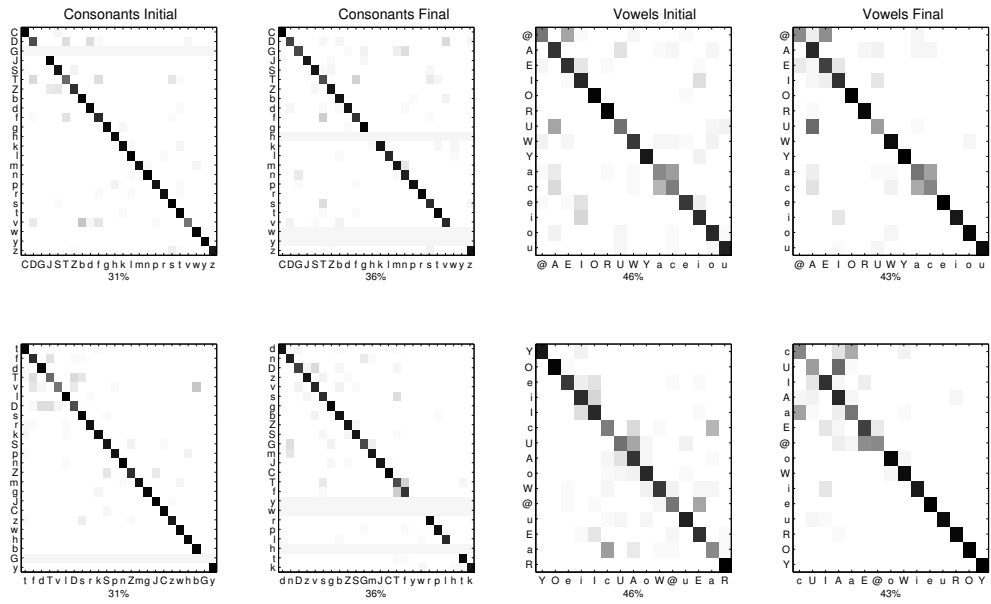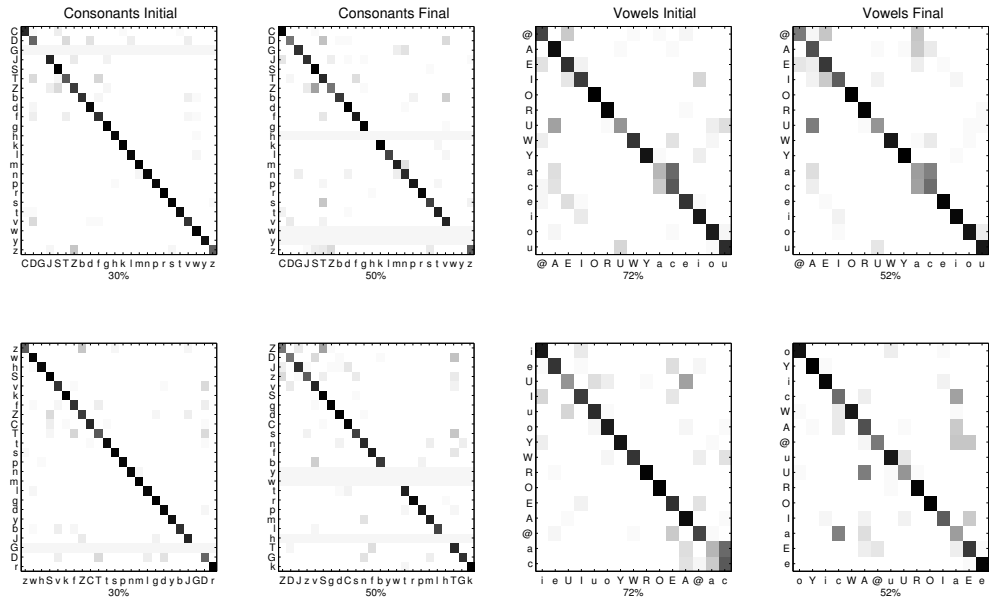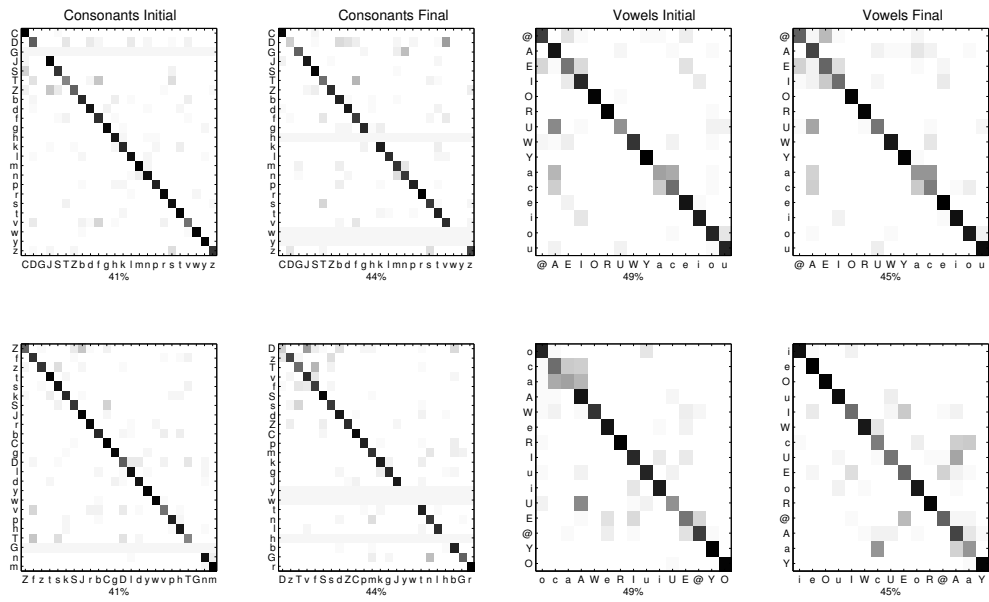
(f) RC-Miguel

Figure F.2: cont.

Appendix G

NSCM CONFUSION PATTERNS AS STACKED BAR PLOTS

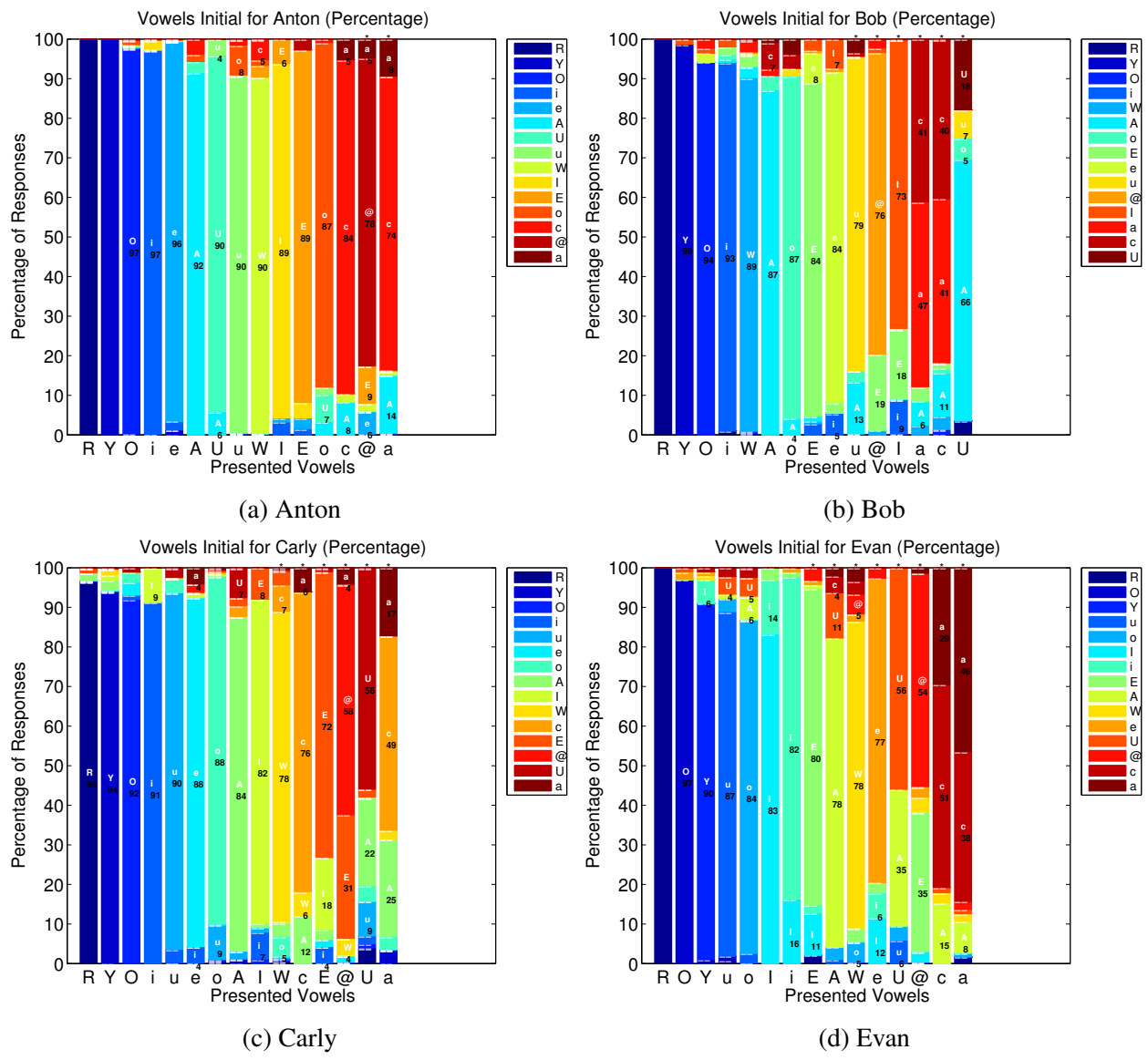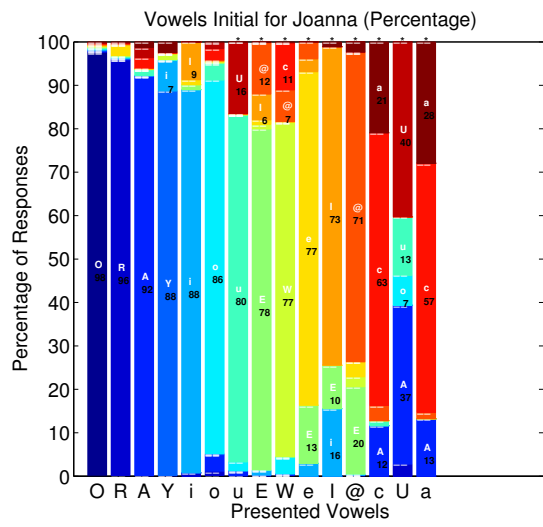(a) Anton
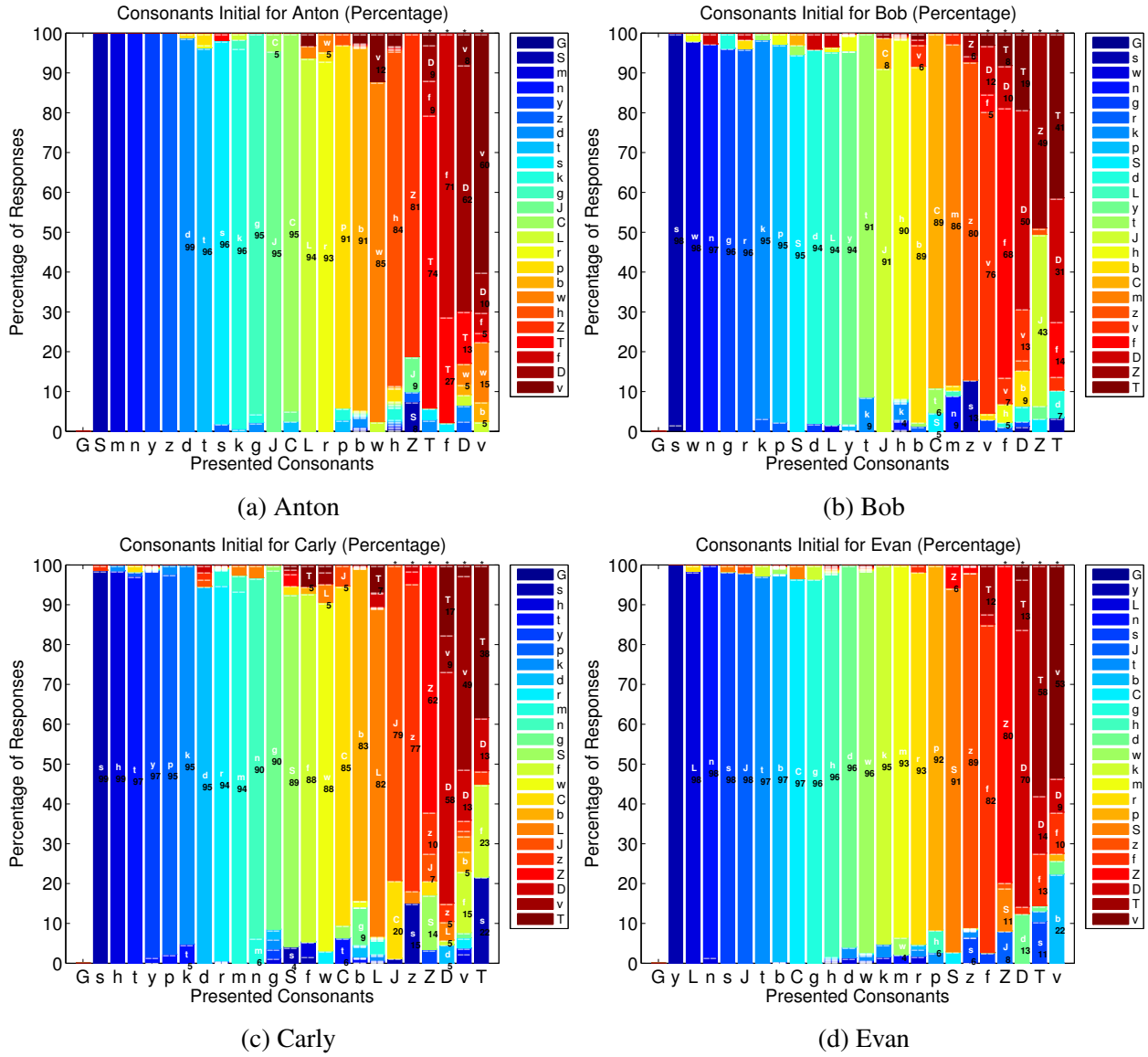
(b) Bob

(c) Carly

(d) Evan

Figure G.1: Stackbar Plots for RCs in CI

(e) Joanna

(f) Miguel

Figure G.1: cont.
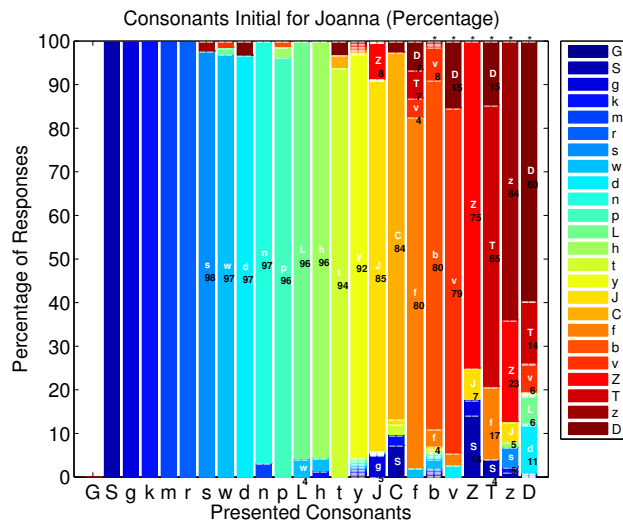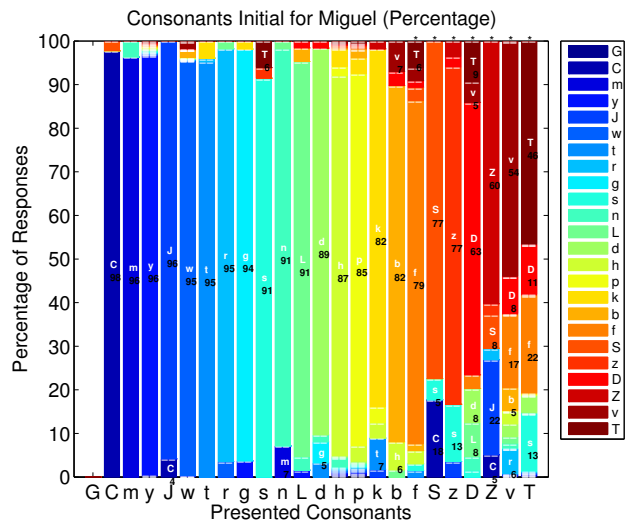
(a) Alina

(b) Angela

(c) Edward
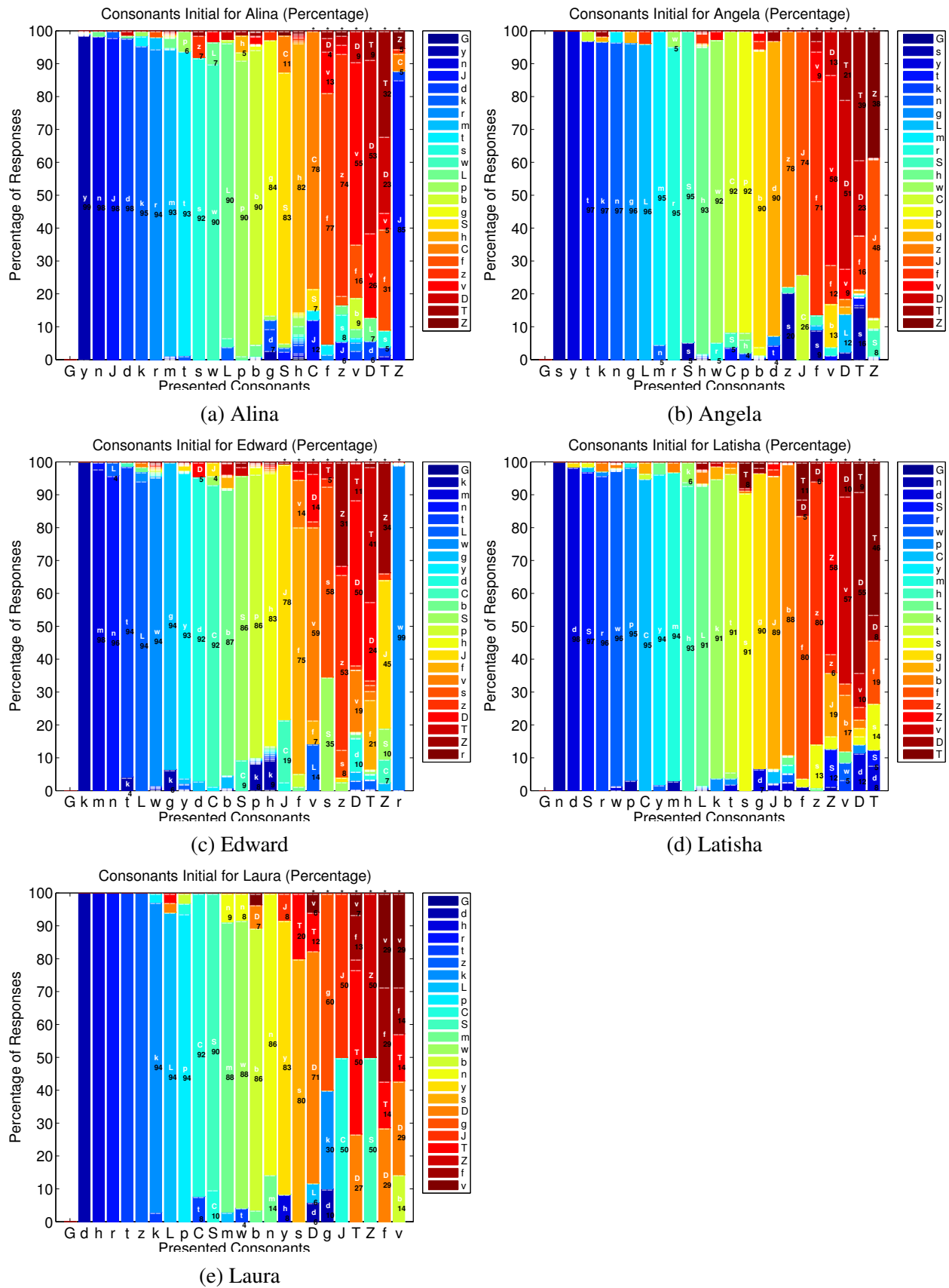
(d) Latisha

(e) Laura

Figure G.2: Stackbar Plots for RDs in CI
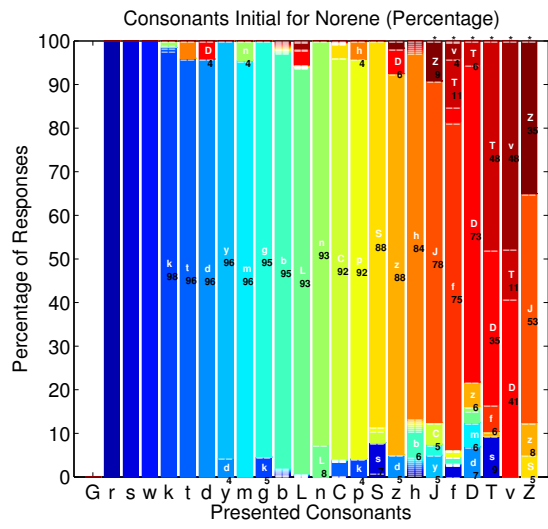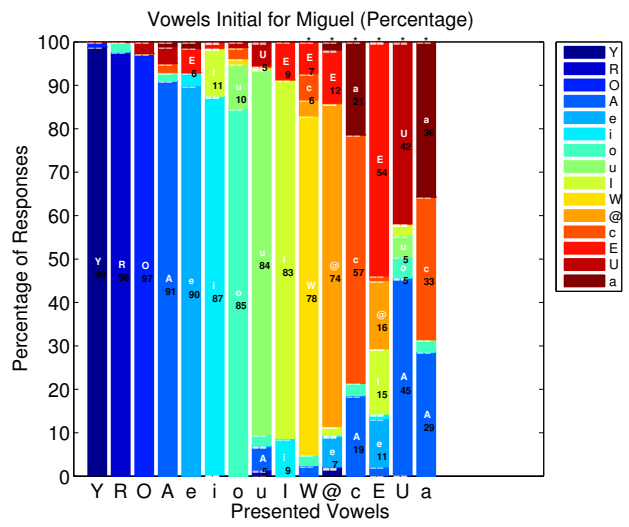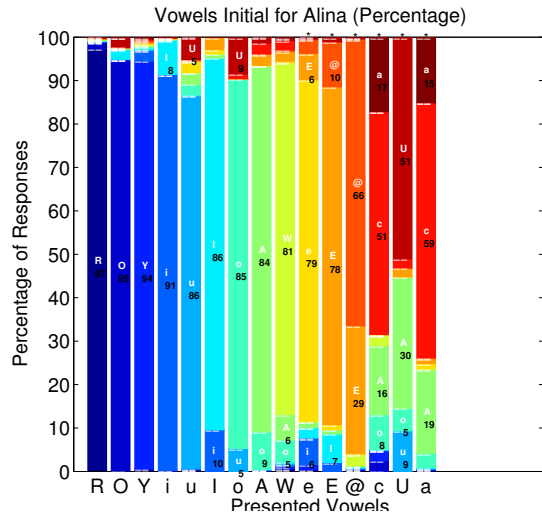
(f) Norene
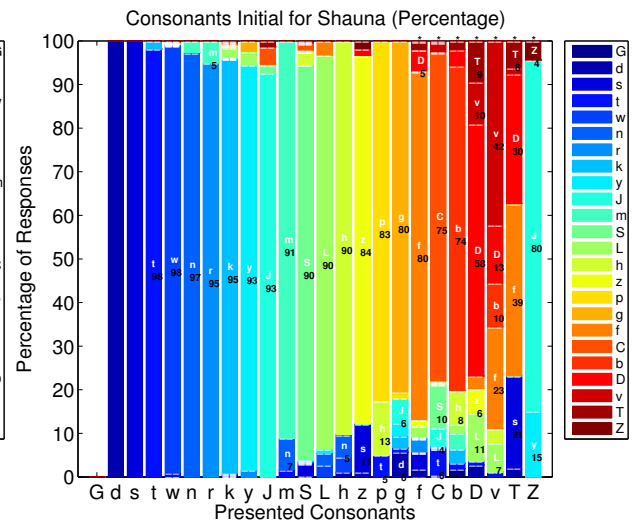
(g) Shauna
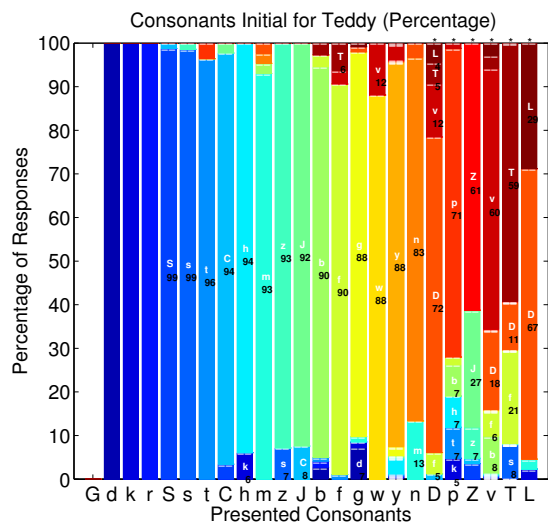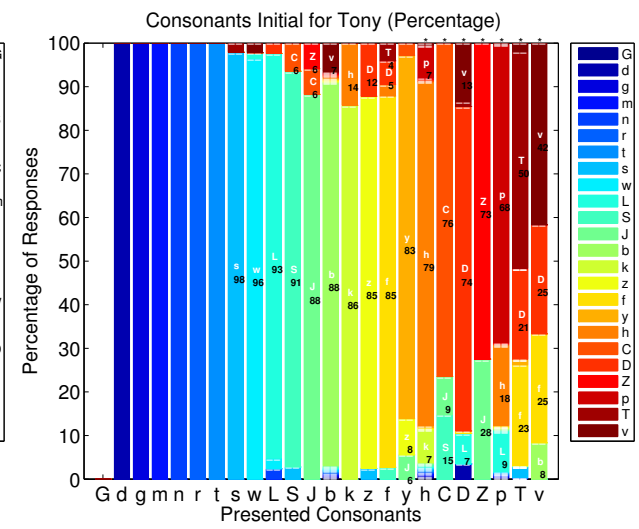
(h) Teddy

(i) Tony

Figure G.2: cont.

165
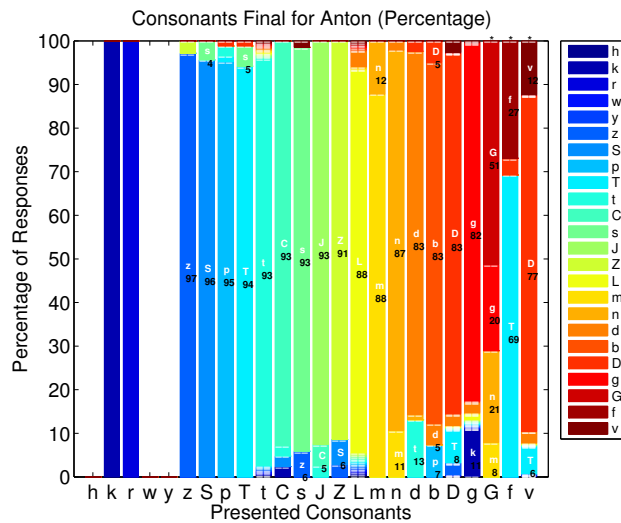
(a) Anton
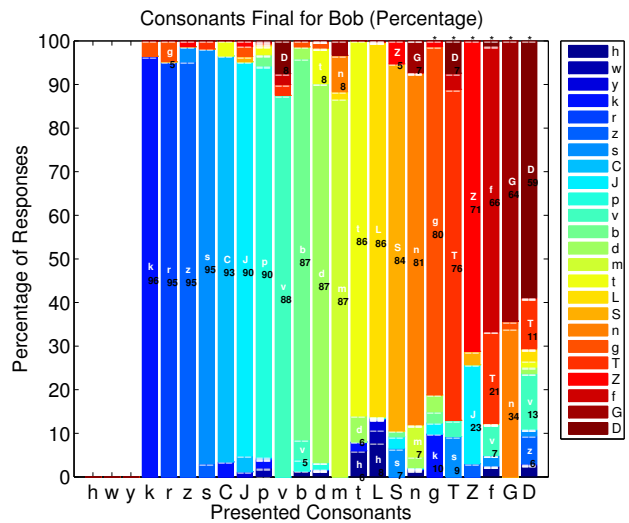
(b) Bob
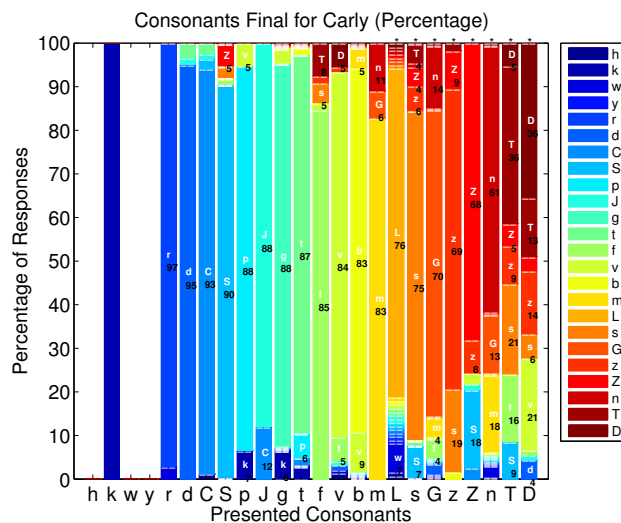
(c) Carly
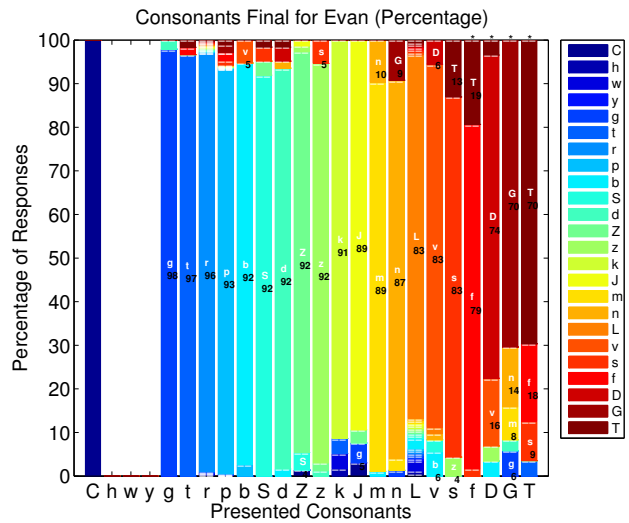
(d) Evan

(e) cont.
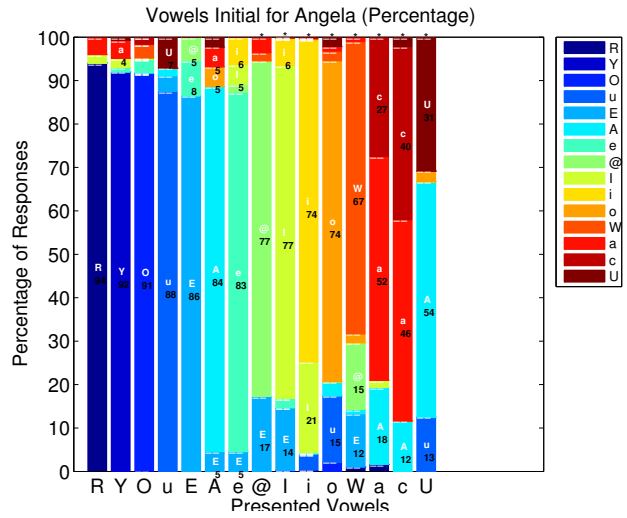
Figure G.3: Stackbar Plots for RCs in CF

(f) Joanna

(g) Miguel

Figure G.3: cont.
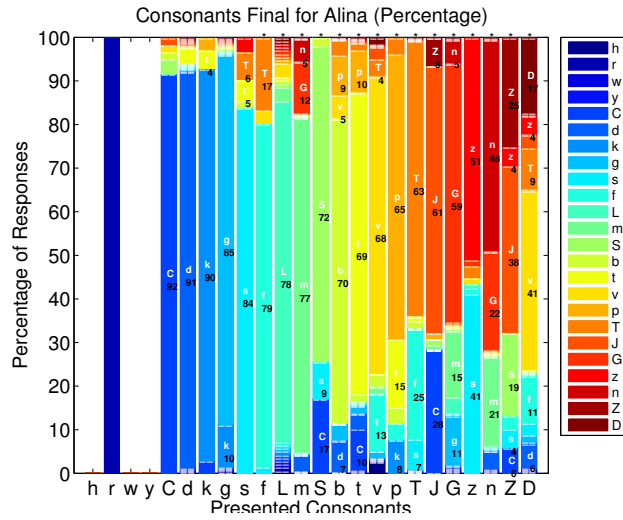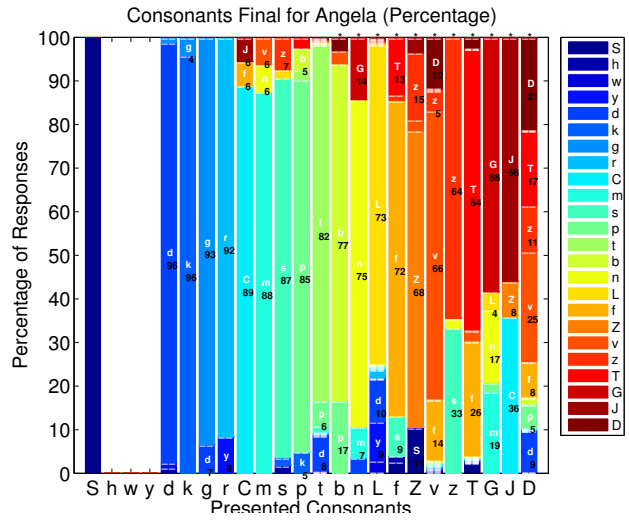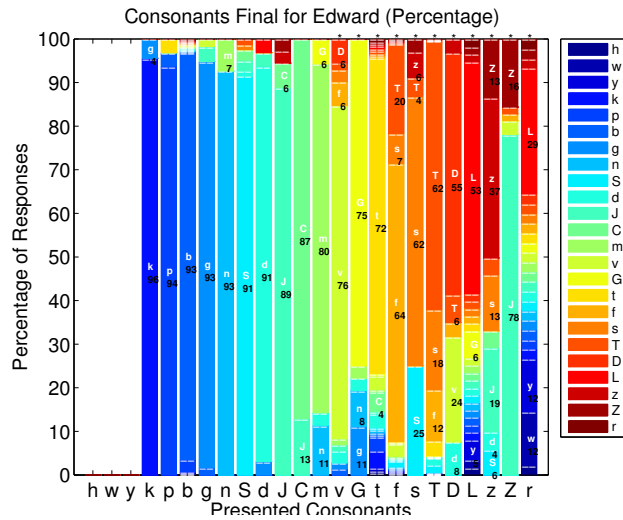
(a) Alina
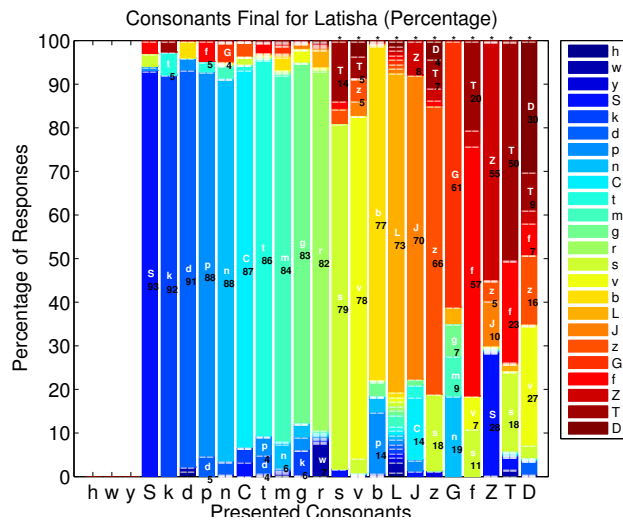
(b) Angela
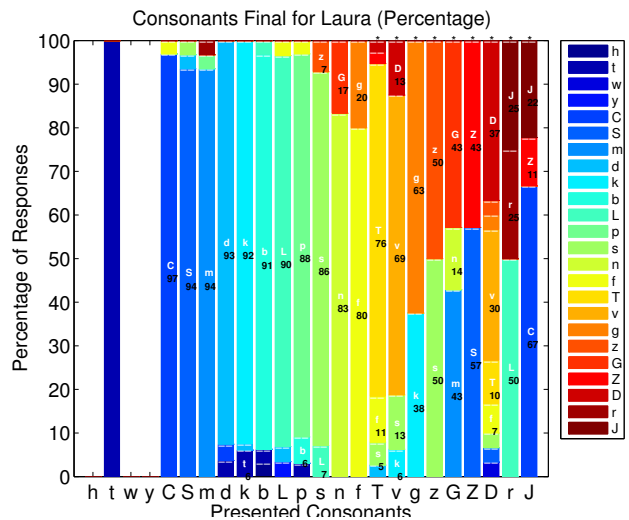
(c) Edward

(d) Latisha

(e) Laura

Figure G.4: Stackbar Plots for RDs in CF

168
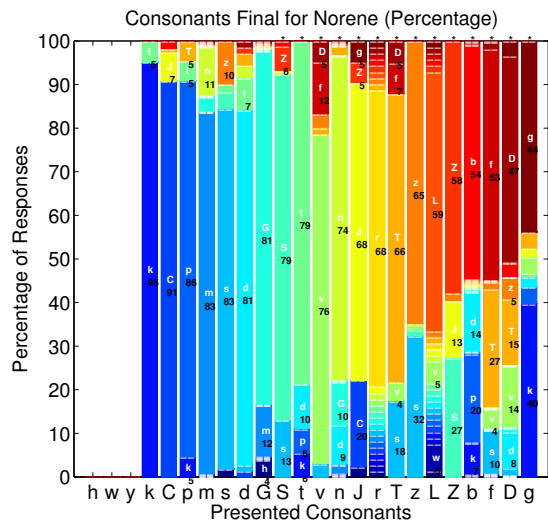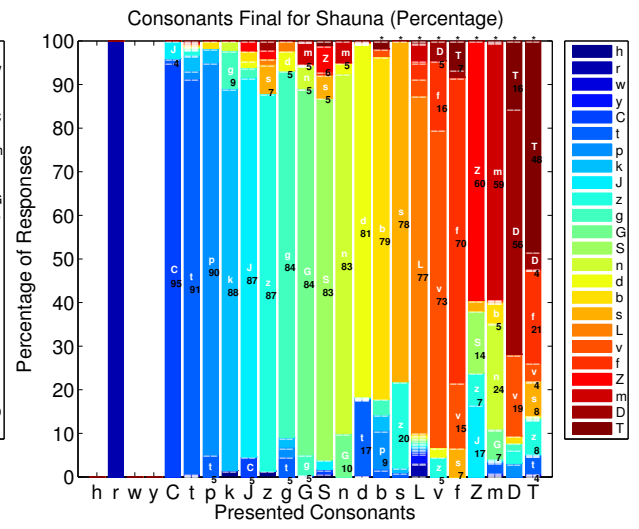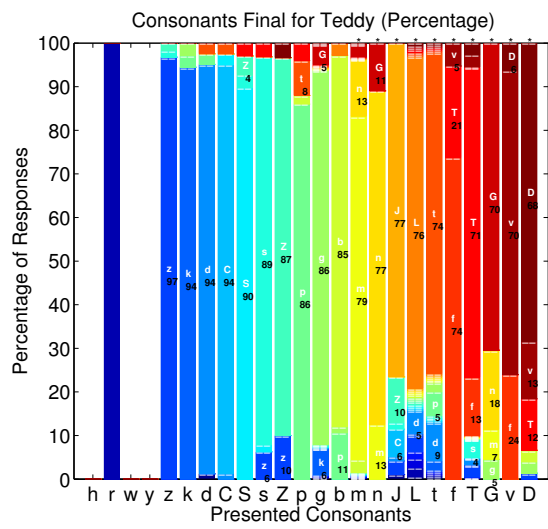
(f) Norene
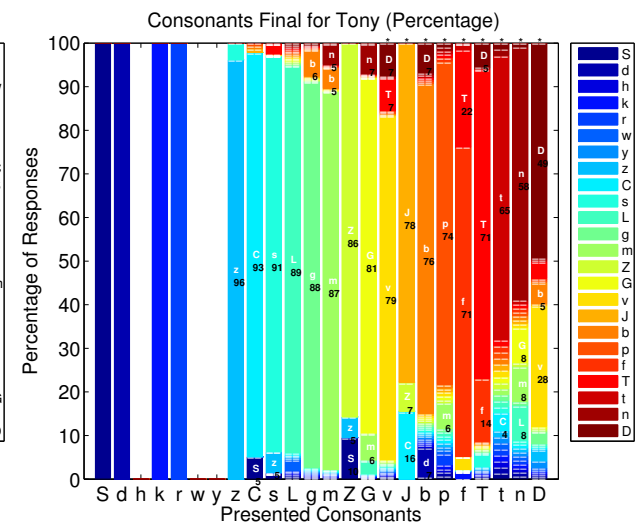
(g) Shauna

(h) Teddy

(i) Tony

Figure G.4: cont.

169

(a) Anton

(b) Bob

(c) Carly

(d) Evan

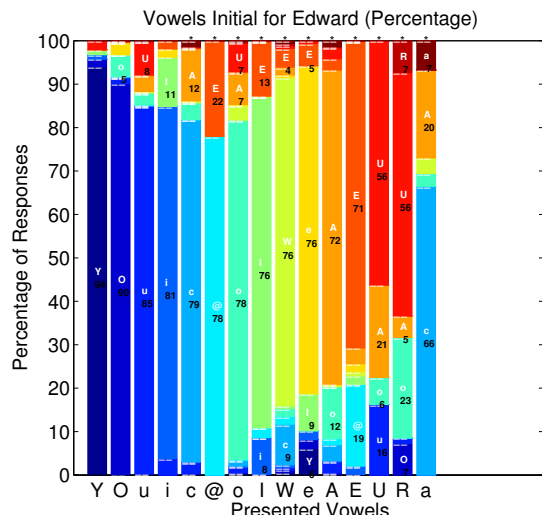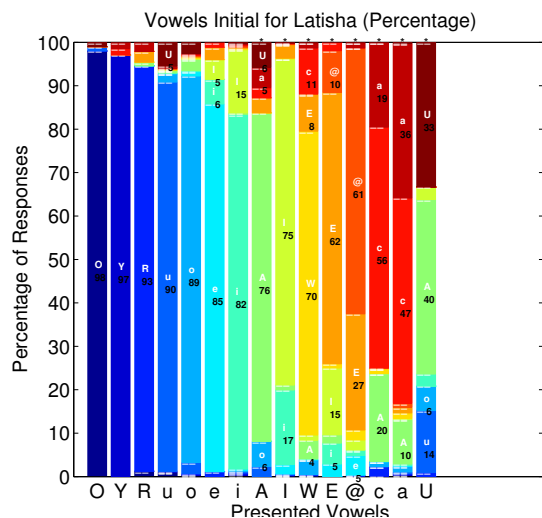Figure G.5: Stackbar Plots for RCs in VI

(e) Joanna (f) Miguel

Figure G.5: cont.

(a) Alina

(b) Angela

(c) Edward

(d) Latisha

(e) Laura

Figure G.6: Stackbar Plots for RDs in VI

(f) Norene

(g) Shauna

(h) Teddy

(i) Tony

Figure G.6: cont.
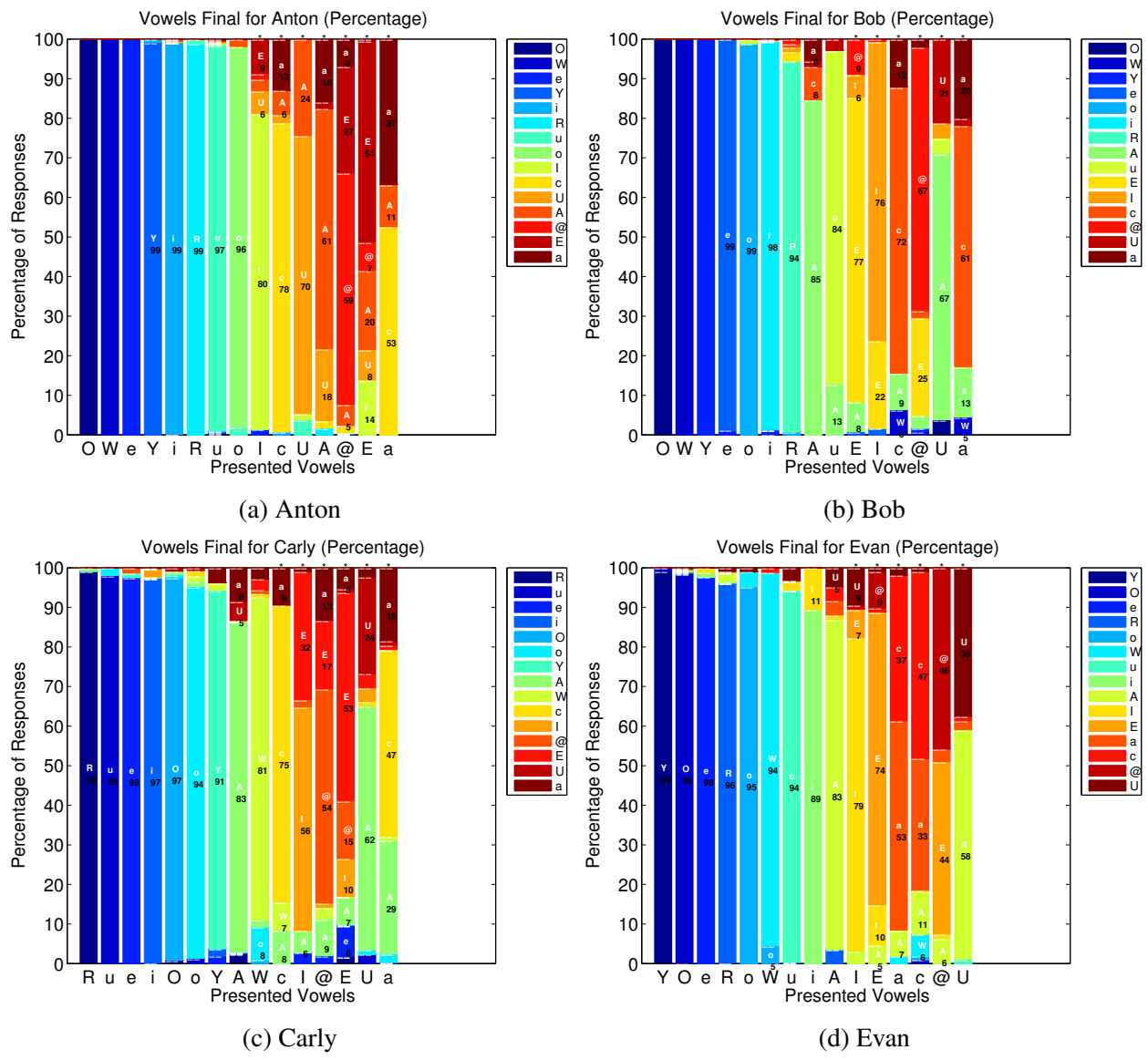
Figure G.7: Stackbar Plots for RCs in VF
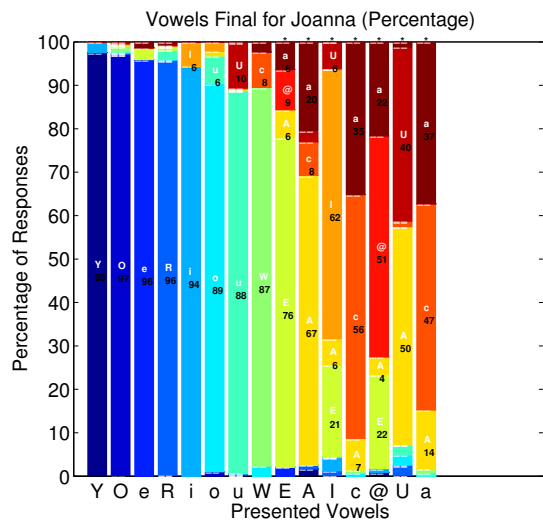
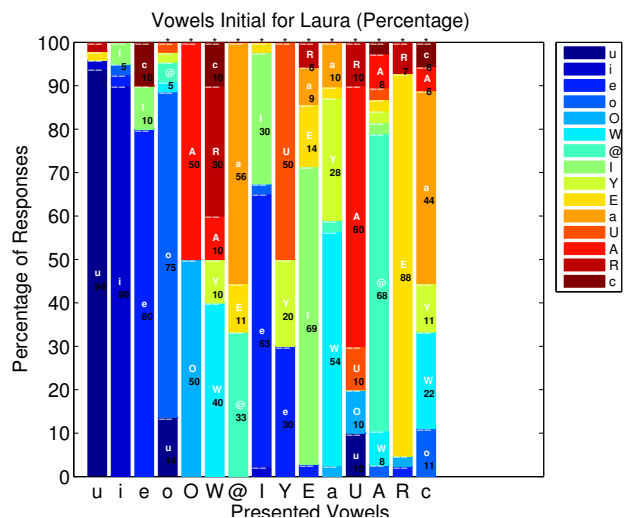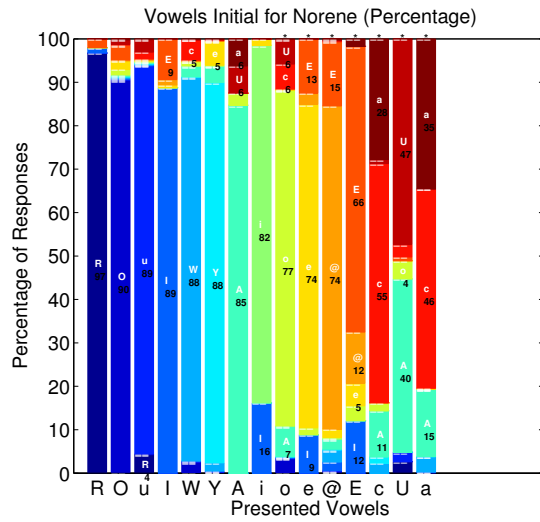(e) Joanna
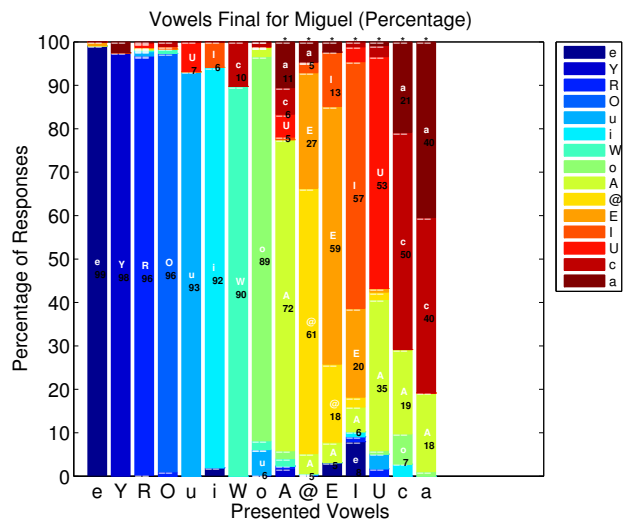
(f) Miguel
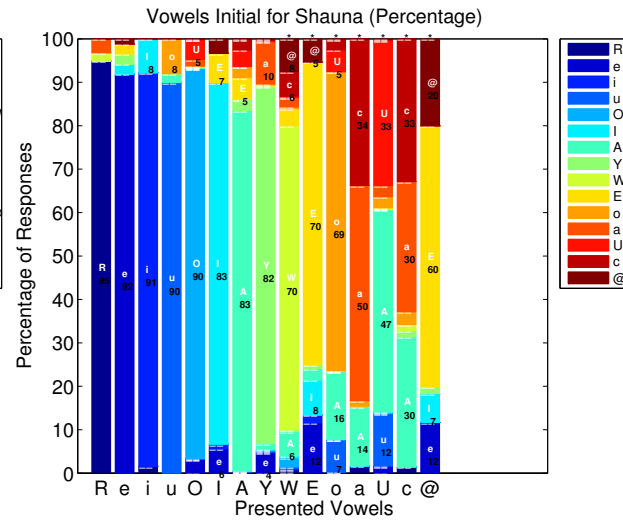
Figure G.7: cont.
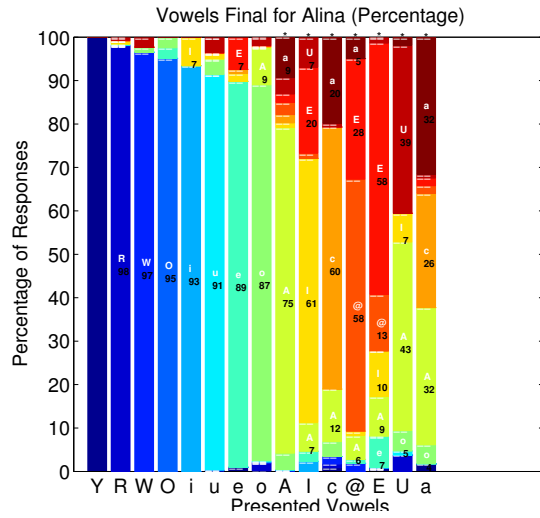
(a) Alina
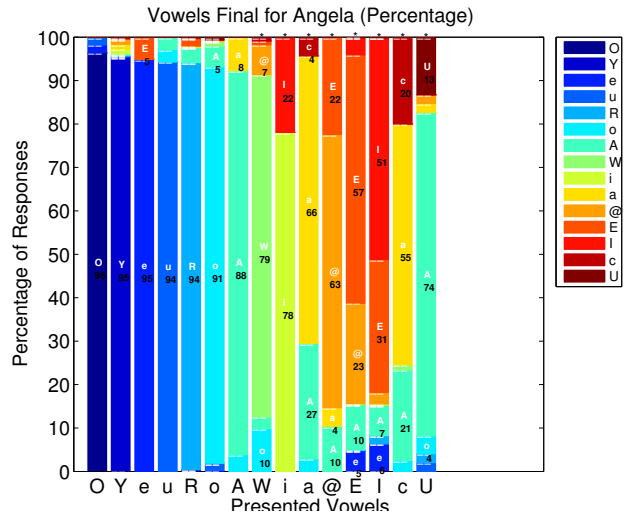
(b) Angela
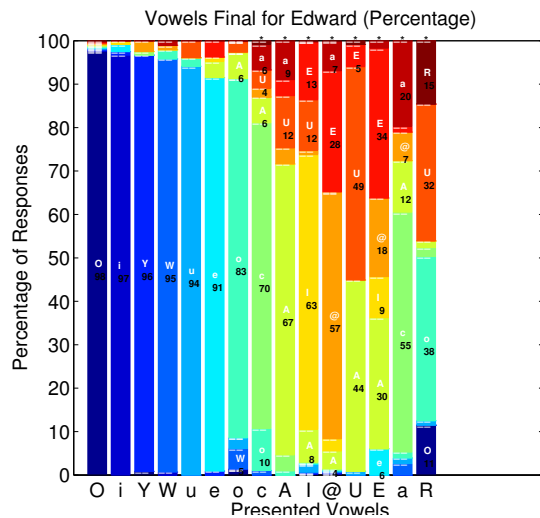
(c) Edward
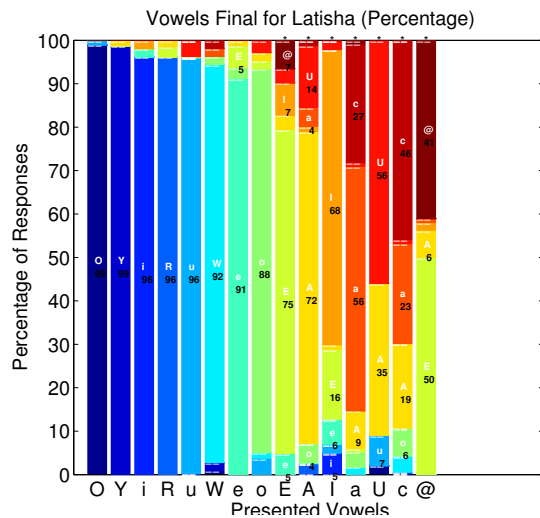
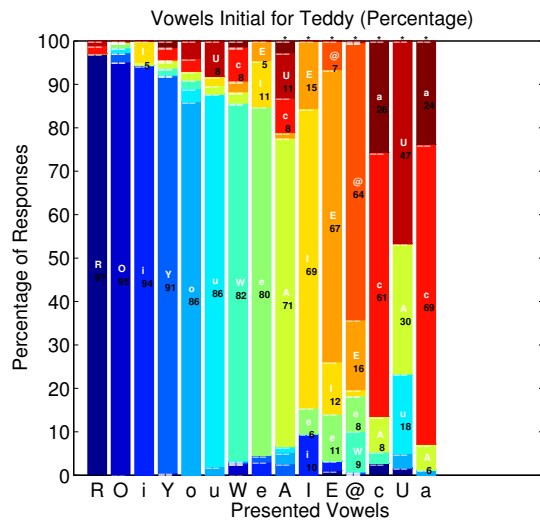(d) Latisha
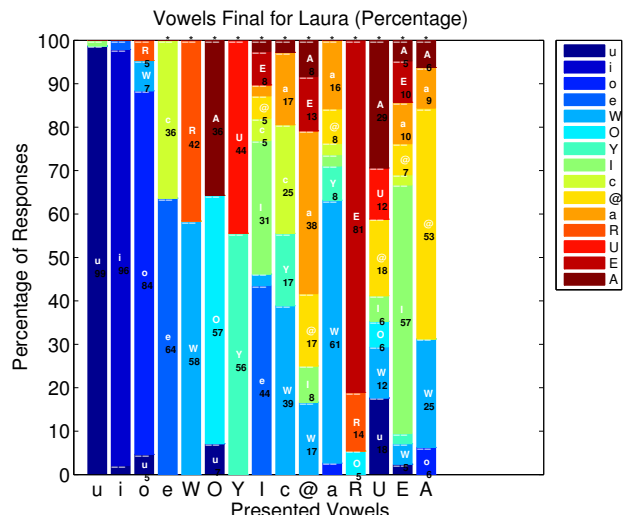
(e) Laura

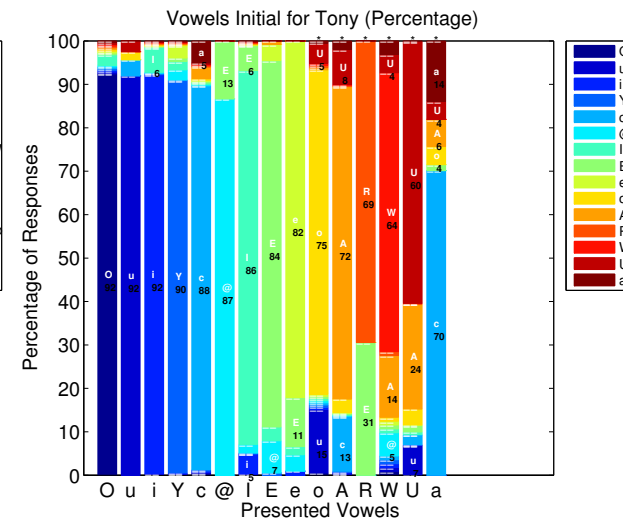Figure G.8: Stackbar Plots for RDs in VF
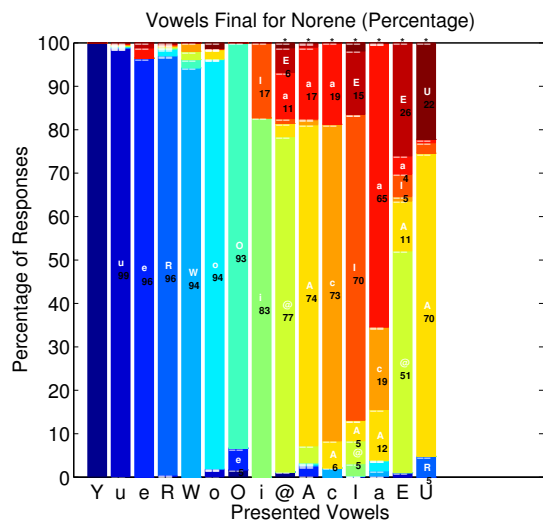
(f) Norene
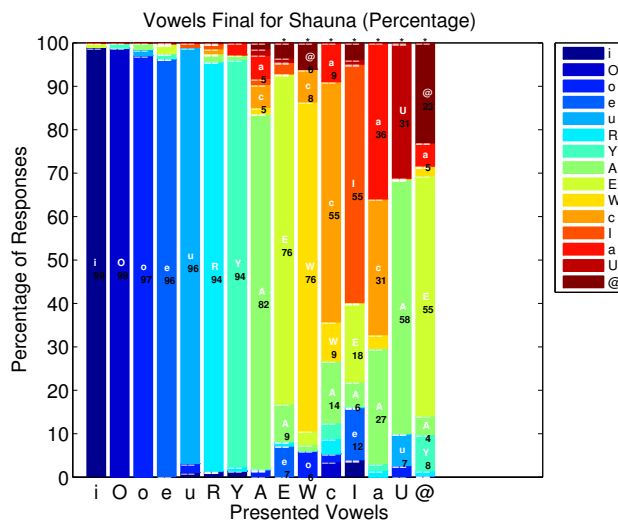
(g) Shauna
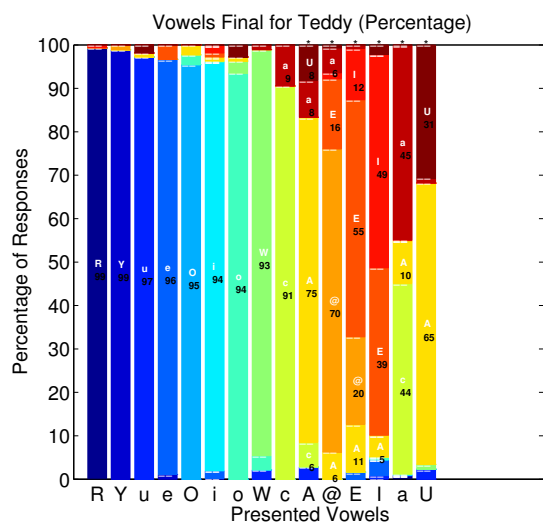
(h) Teddy
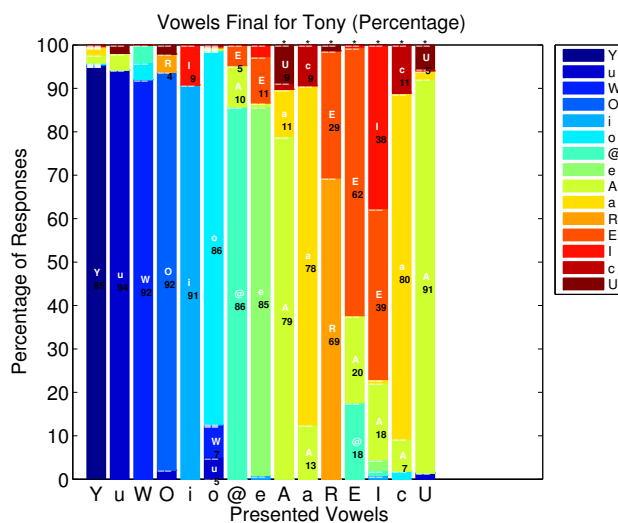
(i) Tony

Figure G.8: cont.

RC PERFORMANCE IN SORTED ERROR PLOTS
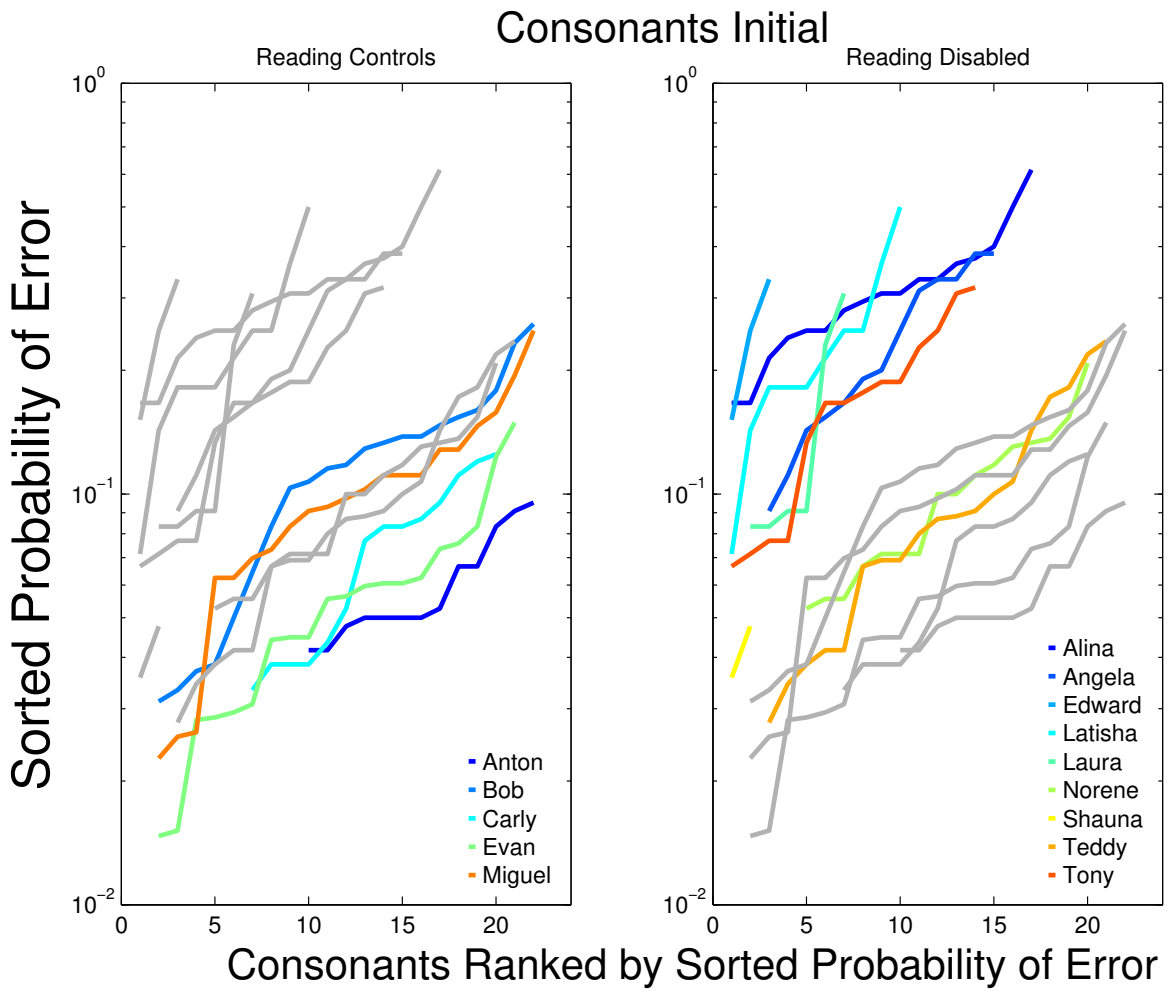
Figure H.1: SCO CI.
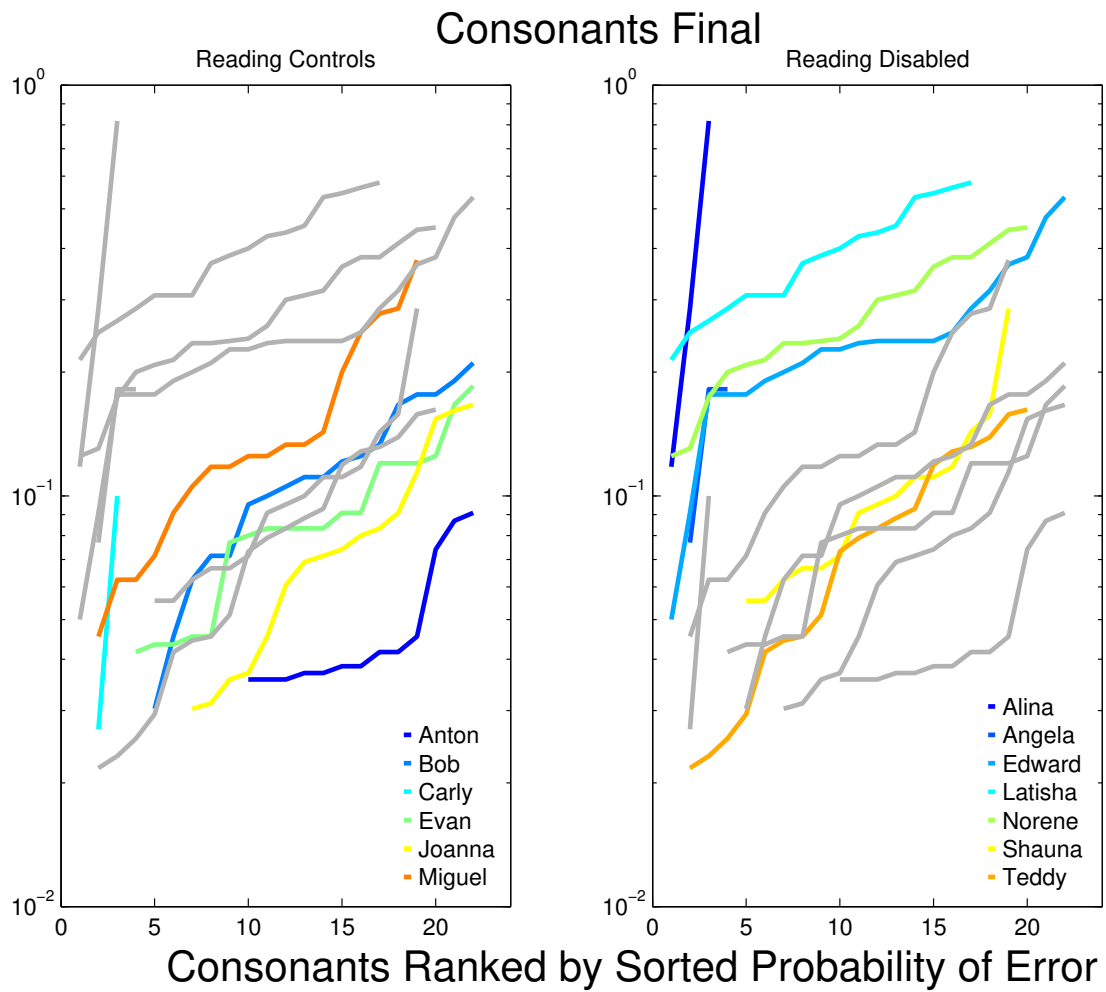
Figure H.2: SCO CF.
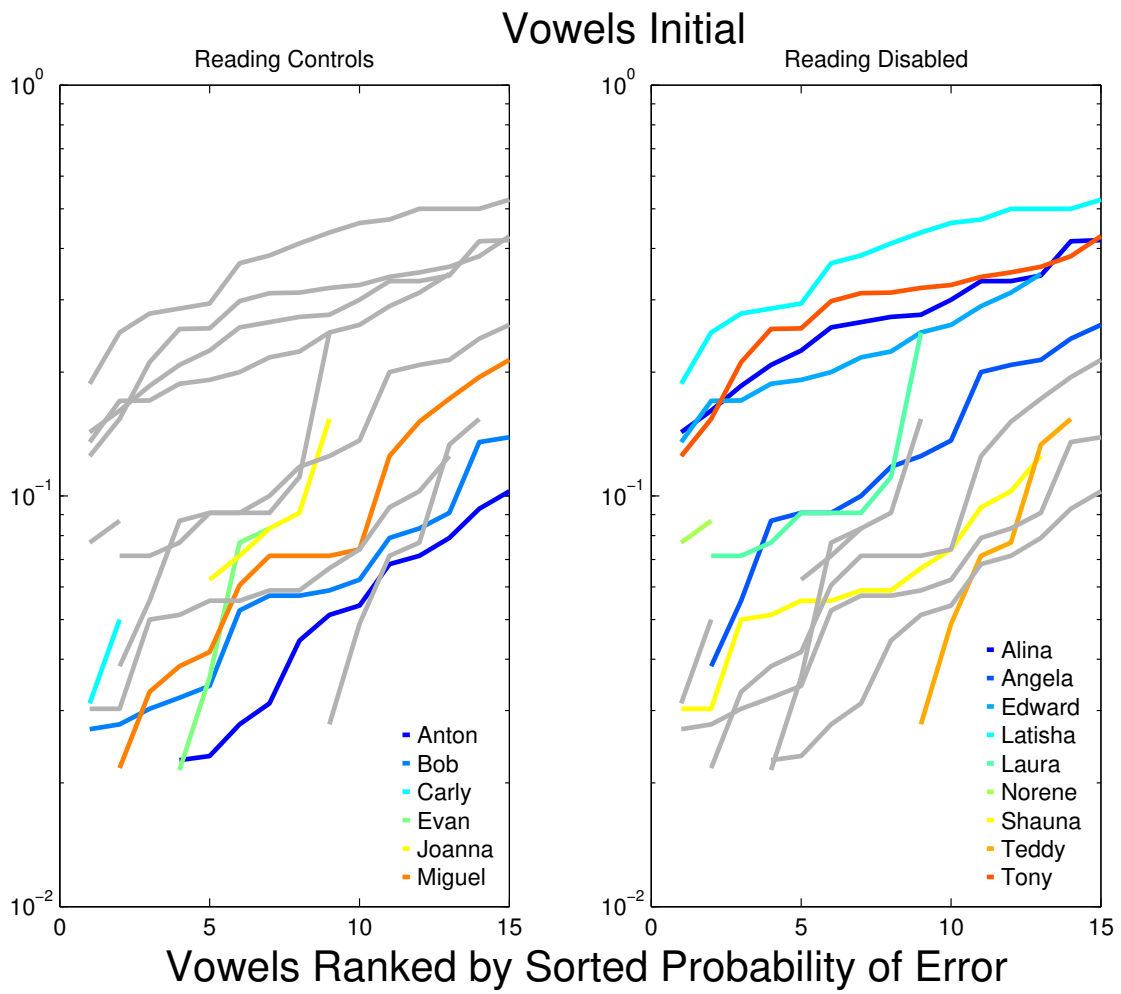
Figure H.3: SCO VI.
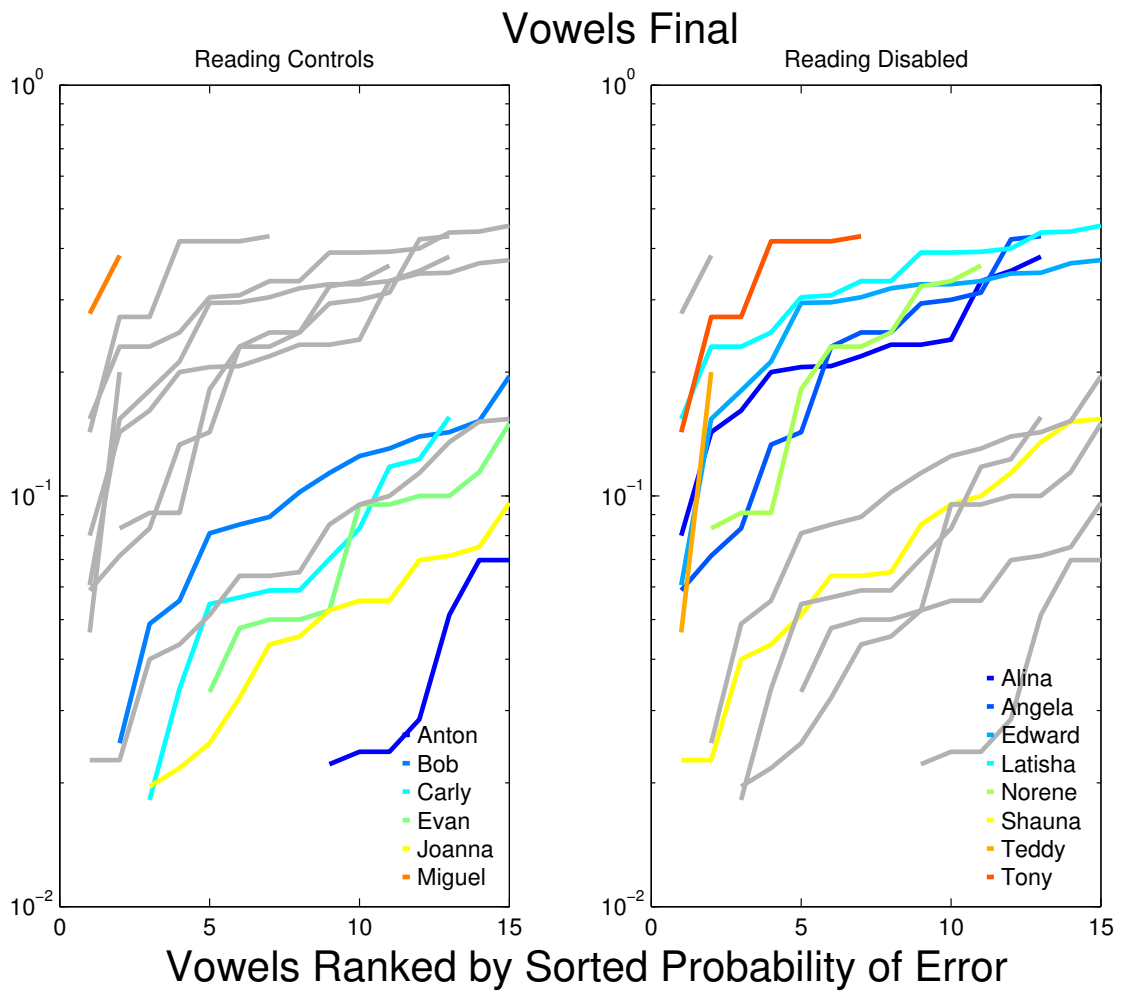
Figure H.4: SCO VF.
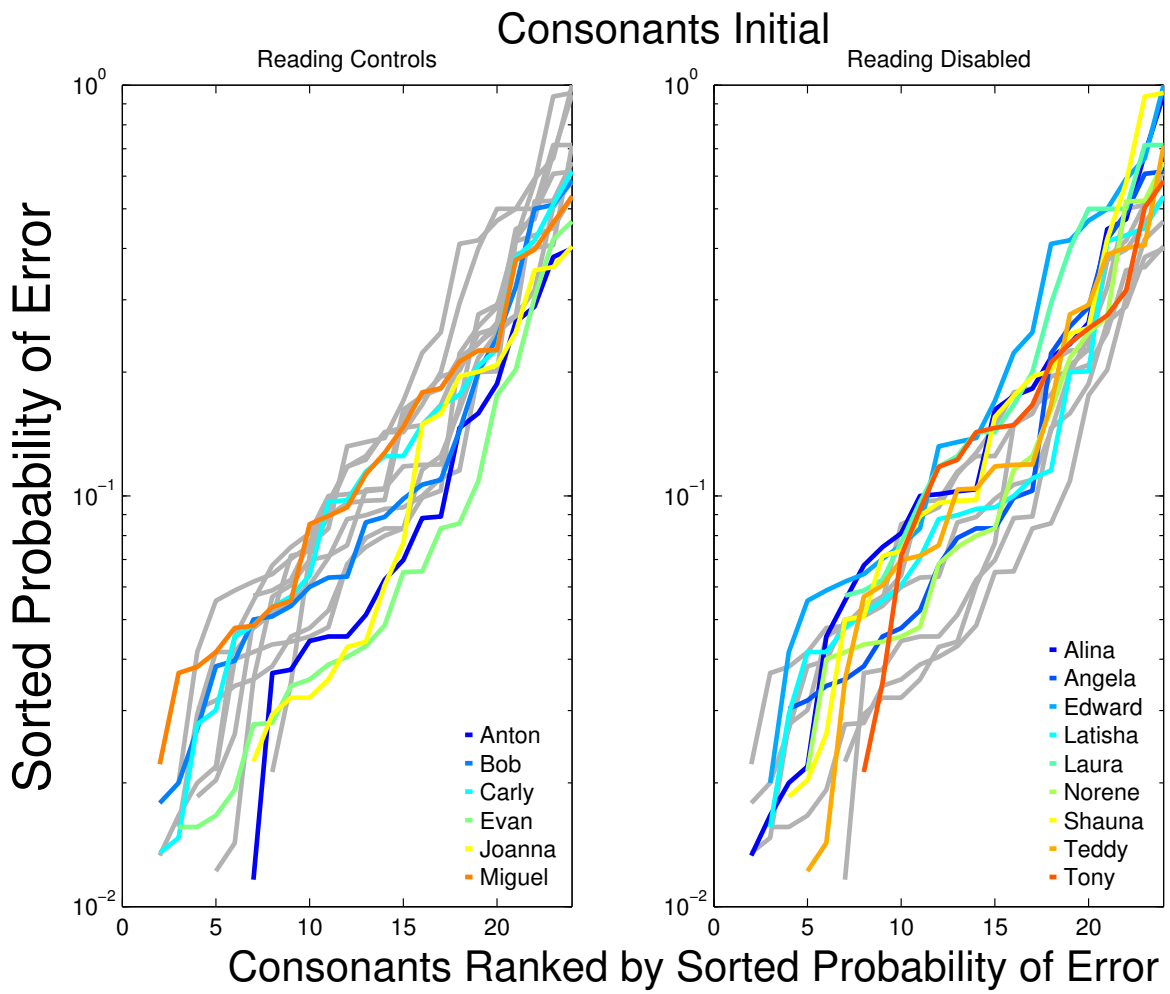
Figure H.5: NSCM CI.
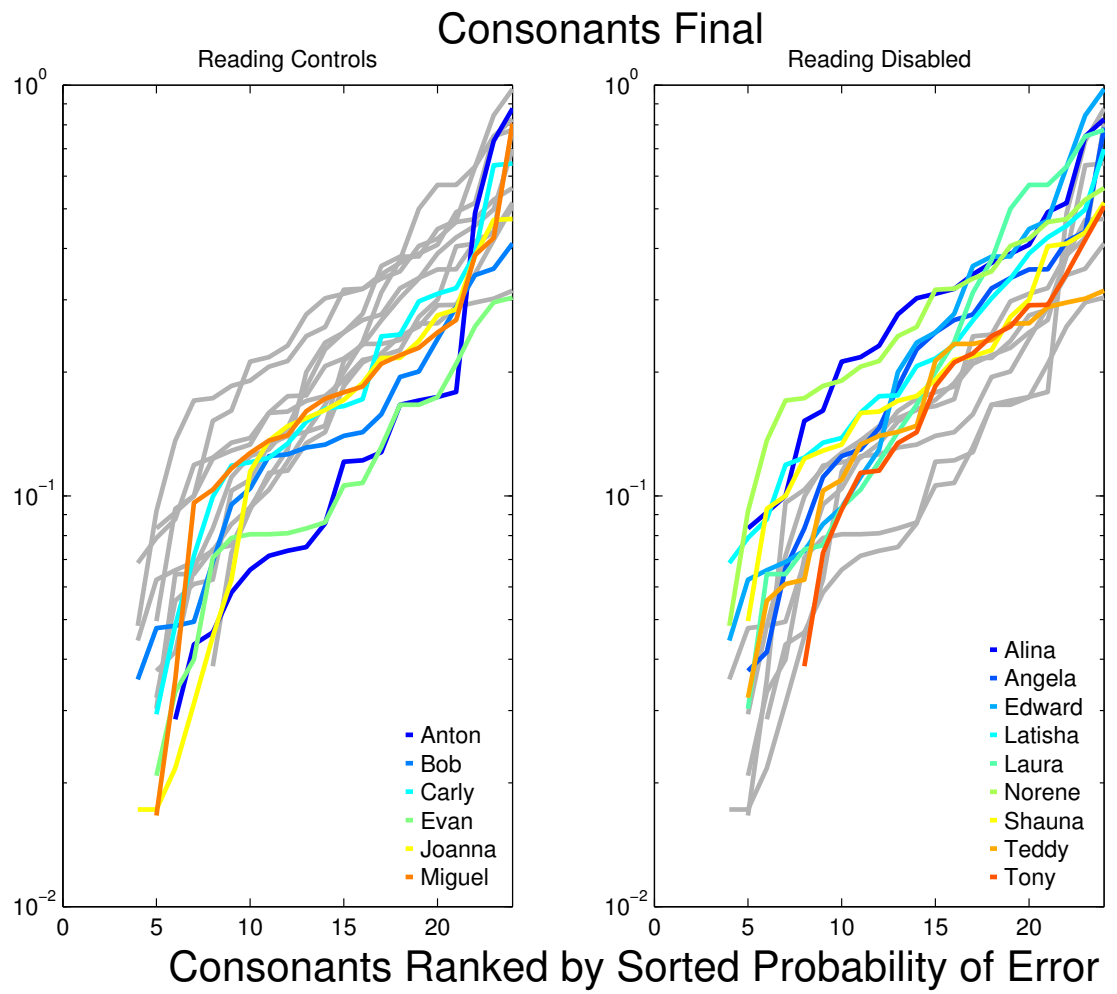
Figure H.6: NSCM CF.
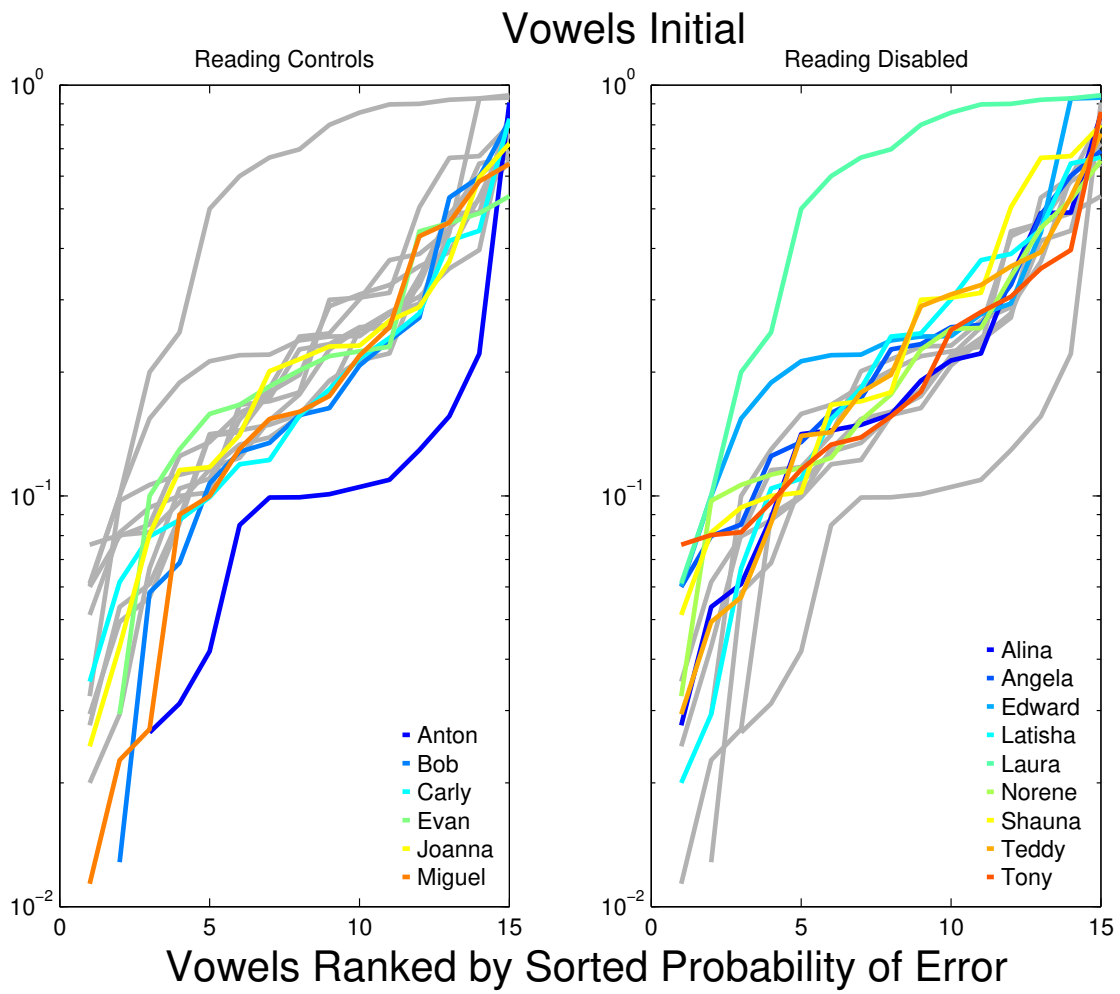
Figure H.7: NSCM VI.
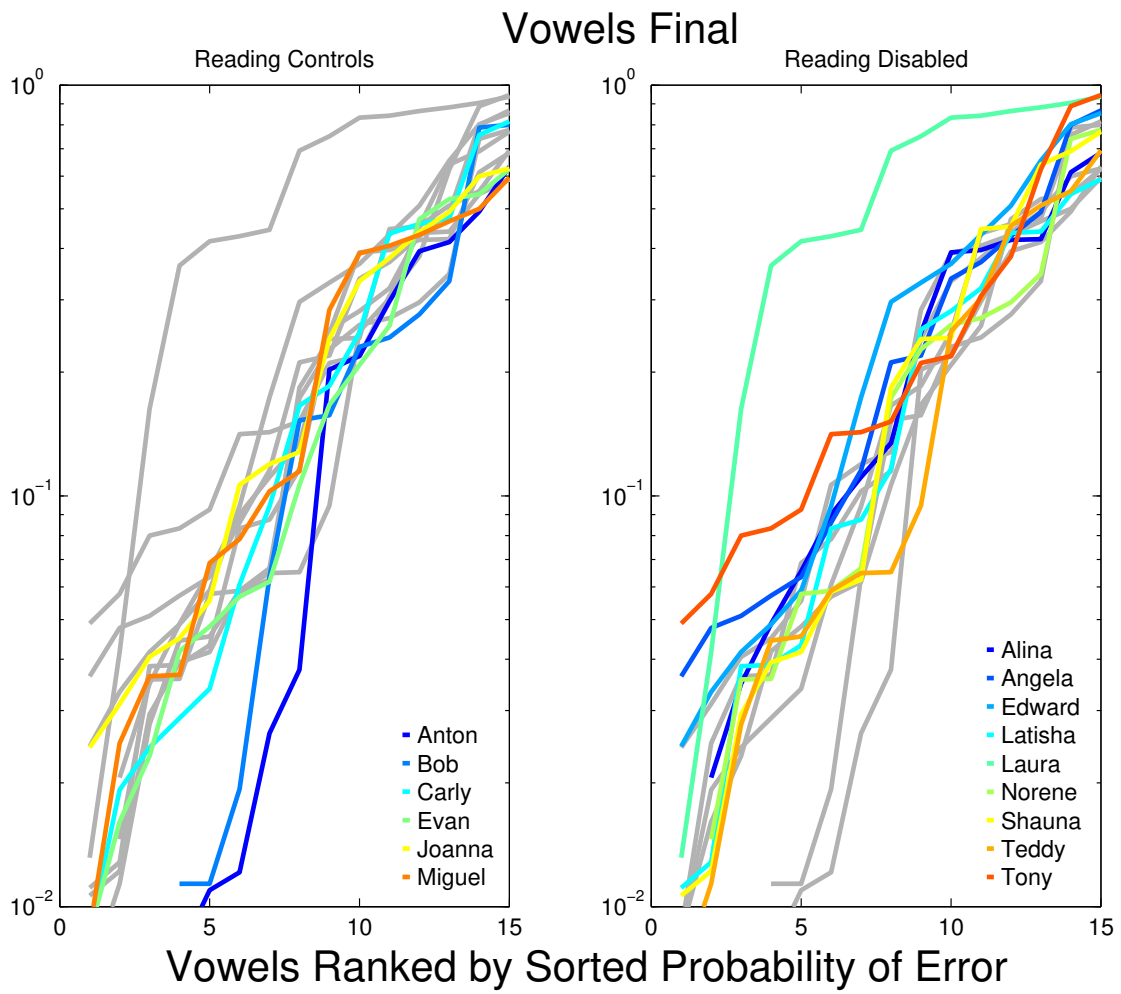
Figure H.8: NSCM VF.
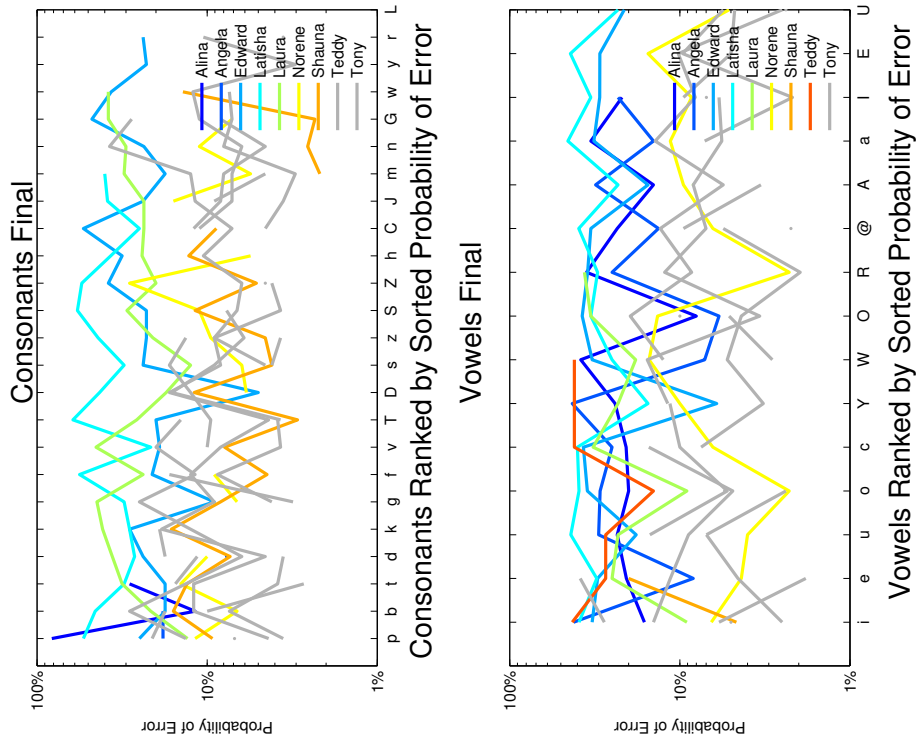
Figure H.9: Unsorted Error Plots for SCO.
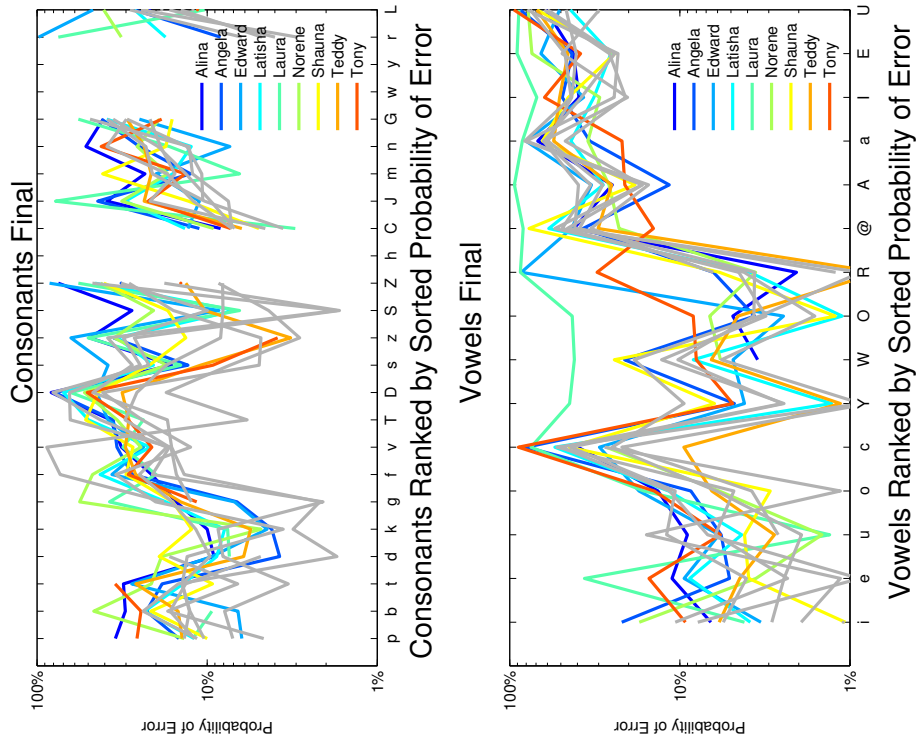
Figure H.10: Unsorted Error Plots for NSCM.

Appendix I

ZOOM IN FOR VI CLUSTER
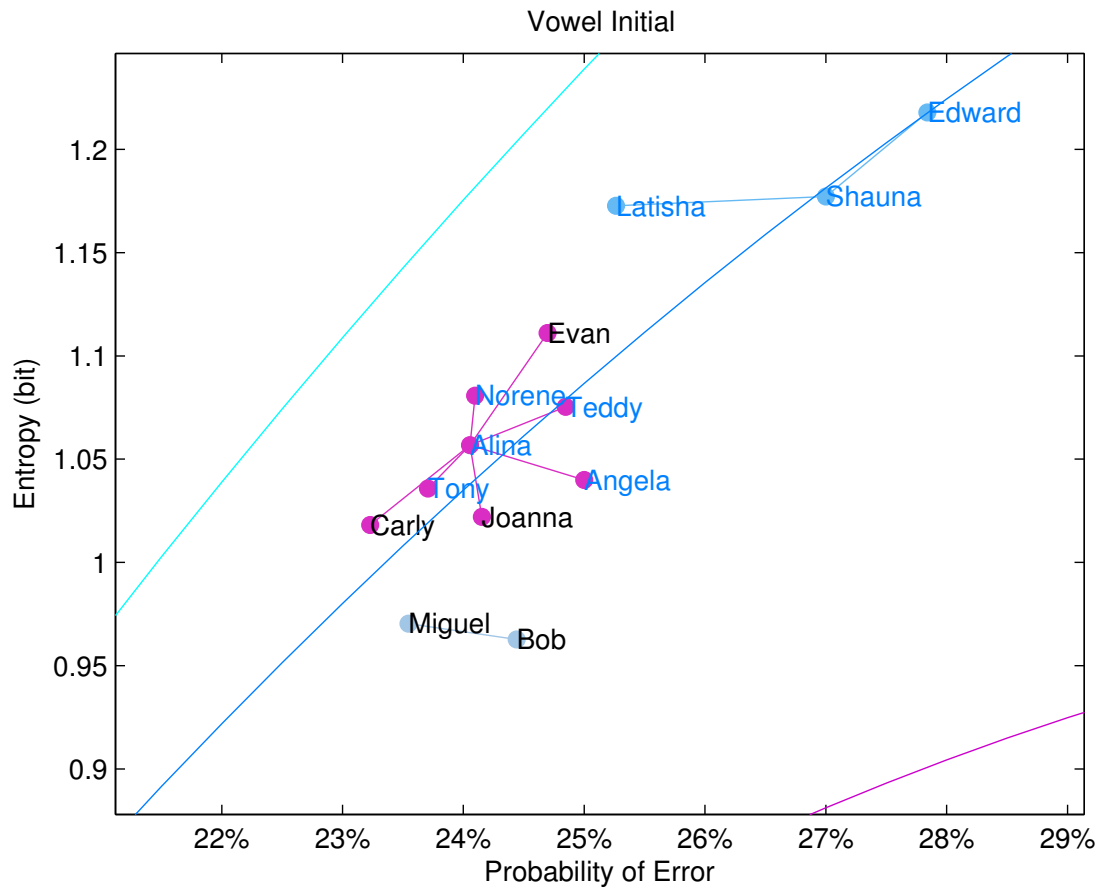
Figure I.1: Central Clusters from Vowel Initial situation of NSCM task.

Appendix J

ANALYSIS PROGRAM

This program `sco.ver7.R` was used to generate figures and do the basic analysis of the SCO data.

```
############################
# Set up environment and list of subjects
############################

# Clear workspace
rm(list = ls(all = TRUE))

# Set minN (number of trials needed per subject per condition to be included in analyses.
# Should be either 1 or 10.
minN = 10

# Load libraries
library(lme4)
library(plyr)
library(ggplot2)

# Set working directory
setwd("~/Desktop/Current/Experiments/RD/ToJoeTosc-Oct19.17/data")

# List of subjects
# Matt Tina Lucas and Savannah were not included in the plot Dec_1 2017 Jie
subjects = c('Alina','Angela','Anton','Bob','Carly','Edward','Evan','Joanna','Latisha',
'Laura','Miguel','Norene','Shauna','Teddy','Tony','Matt','Tina')

# List of RD vs. TD subjects. If subject is in the list below, mark as RD
rd <- ifelse(subjects %in% c('Alina','Angela','Edward','Latisha','Laura','Norene','Teddy',
'Tony','Matt','Tina'),1,0)

# Combine subject and RD list
subRD <- as.data.frame(cbind(subjects,rd))

# Make subject a factor (like a cell in Matlab)
subRD$subject <- as.factor(subRD$subjects)

# Drop old 'subjects' column from table
subRD$subjects <- NULL


############################
```

```
# Read and format data
#############################

# Create an empty data frame (i.e. table) for reading in the data from the Matlab output
# The following columns are in the table:
# phon Which phoneme
# phonType Consonant vs. vowel
# nerr_p_ini Number of errors for P initial position
# n_p_ini Number of trials for P initial position
# pe_ini Proportion of errors for P initial position
# nerr_p_fnl Number of errors for P final position
# n_p_fnl Number of trials for P final position
# pe_fnl Proportion of errors for P final position
# subject Subject ID
scoData <- data.frame(phon=as.character(), phonType=as.character(),
nerr_p_ini=as.numeric(), n_p_ini=as.numeric(),
pe_ini=as.numeric(), nerr_p_fnl=as.numeric(), n_p_fnl=as.numeric(), pe_fnl=as.numeric(),
subject=as.character(), stringsAsFactors=F)

# Loop through list of subjects; read data for consonants
for (s in subjects) {

# Current subject's data. Read from .txt file
current <- as.data.frame(read.delim(paste0('cons/cons_',s,'.txt'),
sep='', comment.char="", header=F, skip=1))

# Add column headers
current <- rename(current, replace=c('V1'='phon', 'V2'='nerr_p_ini', 'V3'='n_p_ini',
'V4'='pe_ini', 'V5'='nerr_p_fnl', 'V6'='n_p_fnl', 'V7'='pe_fnl'))

# Add subject ID
current$subject <- s

# Add phonType
current$phonType <- 'cons'

# Add current subject's data to data frame for all subjects
scoData <- rbind(scoData,current)

}

# Loop through list of subjects; read data for vowels
for (s in subjects) {

# Current subject's data. Read from .txt file
current <- as.data.frame(read.delim(paste0('vows/vows_',s,'.txt'),
sep='', comment.char="", header=F, skip=1))

# Add column headers
current <- rename(current, replace=c('V1'='phon', 'V2'='nerr_p_ini', 'V3'='n_p_ini',
'V4'='pe_ini', 'V5'='nerr_p_fnl', 'V6'='n_p_fnl', 'V7'='pe_fnl'))

# Add subject ID
current$subject <- s
```

```
# Add phonType
current$phonType <- 'vows'

# Add current subject's data to data frame for all subjects
scoData <- rbind(scoData,current)


}

# Convert subject and consonant to factors
scoData$subject <- as.factor(scoData$subject)
scoData$phon <- as.factor(scoData$phon)
scoData$phonType <- as.factor(scoData$phonType)

# Create a new data frame in the 'long' format that R needs to run the analyses. This
# involves converting the pe to trial-level data (error vs. no error for each trial)
# The following columns are defined in the table:
# subject Subject ID
# cons Consonant
# initial Whether the consonant was in initial position (1) or final (0)
# correct Whether the subject made a correct response or not (1 vs. 0)
newData <- data.frame(subject=as.character(), phon=as.character(), phonType=as.numeric(),
initial=as.numeric(), correct=as.numeric(), stringsAsFactors=F)

# Loop through each row in consData (i.e. the data frame created above from reading in the
# .txt files
for (i in 1:nrow(scoData)) {

# Get the current row
current <- scoData[i,]

# First, convert the P initial errors into trial-level table
# Get number of P initial trials
n_ini <- current$n_p_ini
# Get number of errors
n_err_ini <- current$nerr_p_ini
# Create an empty data frame with the same columns as newData
newCurrent <- newData[NULL,]
# counter variable
k = 1

# Include any condition with at least one trial (n_ini >= 1) or only those with at
# least 10 trials (n_ini >= 10). Set minN at beginning of script to 1 or 10.
if (n_ini >= minN) {

# Loop from 1 to number of trials for this condition
for (j in 1:n_ini) {
# Mark the initial column as 1 (these are P initial trials)
newCurrent[k,]$initial <- 1
# Add the consonant label
newCurrent[k,]$phon <- as.character(current$phon[1])
# Add the subject ID
newCurrent[k,]$subject <- as.character(current$subject[1])
# Add phonType
newCurrent[k,]$phonType <- as.character(current$phonType[1])
# Add whether or not they are correct for that trial. If the current
```

191

```
# trial (j) is less than the number of errors for this condition, mark
# as incorrect; otherwise mark as correct
newCurrent[k,]$correct <- ifelse (j <= n_err_ini, 0, 1)
# Increment the row counter
k = k + 1
}
}


# Add the current data to the table with all data (newData)
newData <- rbind(newData,newCurrent)


# Same as above, but for P final position
n_fnl <- current$n_p_fnl
n_err_fnl <- current$nerr_p_fnl
newCurrent <- newData[NULL,]
k = 1
if (n_fnl >= minN) {
for (j in 1:n_fnl) {
newCurrent[k,]$initial <- 0
newCurrent[k,]$phon <- as.character(current$phon[1])
newCurrent[k,]$subject <- as.character(current$subject[1])
newCurrent[k,]$phonType <- as.character(current$phonType[1])
newCurrent[k,]$correct <- ifelse (j <= n_err_fnl, 0, 1)
k = k + 1
}
}
newData <- rbind(newData,newCurrent)


}


# Add column for error (1 - correct)
newData$error <- 1 - newData$correct

# Make subject and consonant a factor
newData$subject <- as.factor(newData$subject)
newData$phon <- as.factor(newData$phon)

# Create numeric phonType (cons==1, vows==0)
newData$phonType <- ifelse(newData$phonType=='cons',1,0)

# Merge the table with RD status into the data table
newData <- merge(newData,subRD)

# Save a copy of newData (in case it is needed later)
origNewData <- newData



############################
# Stats
############################

# The overall approach is to run a "mixed-effects regression" on the errors. Because error
# is binomially distributed, we use the logit link function (i.e. a logistic regression)
# Because some variables are random (subject and consonant) and others are fixed (RD and
# C_position) it is a "mixed-effects" model. The glmer function (from the lme4 package)
```

192

```
# allows us to fit mixed-effects models of this type. The output will test the significance
# of various main effects and interactions

# Convert RD into a numeric variable
newData$rd <- as.numeric(newData$rd)

# Center RD, initial, and phonType variables; this makes the regression models more interpretable
newData$rd <- newData$rd - mean(newData$rd)
newData$initial <- newData$initial - mean(newData$initial)
newData$phonType <- newData$phonType - mean(newData$phonType)

# For the random effects, statistical significance is tested via model comparison. That is,
# does adding a random effect to the model significantly improve the model fit? If so,
# this is a statistically significant effect. The random effects are added to the model in the
# following order:
#
# subject Main effect of subject (i.e. do subjects differ in their overall error rate?)
# initial x cons Interaction (does the consonant's error rate vary depending on position?)
# initial x cons Interaction (does subject's error rate vary depending on position?)
# sub x cons Interaction (do invidual subjects vary in which consonants cause errors?)
# sub x cons x initial 3-way interaction (do all three factors interact?)
#
# Note: This does not test for the main effect of consonant (do some consonants lead to more errors
# than others?). This is because glmer requires at least one random effect in the initial model.
# I chose to not test for the main effect of consonant, since it seemed to be the least relevant
# and we already know that some consonants produce more errors than others. Altneraitively, a
# parallel analysis could be done that starts with a main effect of subject and then tests for the
# consonant effect.

model1a <- glmer(error ~ initial*rd*phonType +
(1|phon),
data=newData, family='binomial')

model2a <- glmer(error ~ initial*rd*phonType +
(1|phon) + (1|subject),
data=newData, family='binomial')

model3a <- glmer(error ~  initial*rd*phonType +
(1|phon) + (1|subject) + (0+initial|phon),
data=newData, family='binomial')

model4a <- glmer(error ~  initial*rd*phonType +
(1|phon) + (1|subject) + (0+initial|phon) + (0+initial|subject),
data=newData, family='binomial')

model5a <- glmer(error ~  initial*rd*phonType +
(1|phon) + (1|subject) + (0+initial|phon) + (0+initial|subject) + (0+phonType|subject),
data=newData, family='binomial')

model6a <- glmer(error ~ initial*rd*phonType + (1|phon) + (1|subject) + (0+initial|phon) +
(0+initial|subject) + (0+phonType|subject) + (0+initial:rd|phon), data=newData, family='binomial')

model7a <- glmer(error ~ initial*rd*phonType +
(1|phon) + (1|subject) + (0+initial|phon) + (0+initial|subject) + (0+phonType|subject)
+ (0+initial:rd|phon) + (0+initial:phonType|subject), data=newData, family='binomial')
```

```
model8a <- glmer(error ~ initial*rd*phonType +
(1|phon) + (1|subject) + (0+initial|phon) + (0+initial|subject) + (0+phonType|subject)
+ (0+initial:rd|phon) + (0+initial:phonType|subject) + (1|phon:subject), data=newData,
family='binomial')

# After fitting each model, the anova function is used to compare them to test for signifiance.
# Note, this does not run an ANOVA!! It is actually a Chi-square test comparing the goodness-of-
# fit for each model. The output of this is saved to 'randomEffects'.
randomEffects <- anova(model1a,model2a,model3a,model4a,model5a,model6a,model7a,model8a)

# To evaluate the fixed effects (RD and initial), we can simply get the summary of the most
# complex model. This is saved to 'fixedEffects'.
fixedEffects <- summary(model8a)

## ANOVA (for comparison to regression)

# ANOVA analysis with arcsine transformed data. Note: This is not the best approach, but should
# be a useful approximation to compare with regression. See Jaeger (2008), Journal of Memory
# and Language. Note also: This does not test for random effects (i.e., individual differences
# between subjects and consonants).

#asinTransform <- function(q) {46.47324337*2*asin(sqrt(q))-23}
#pAsin <- asinTransform(sco$p)
#sco.aov <-aov(pAsin~ sub*pos+Error(1/individual), data = sco)
#summary(sco.aov)

# ANOVA needs subject averages. The next line computes those.
dataAvg <- ddply(newData, .(subject, initial, rd, phonType), summarize, pe=mean(error))

# Make sure everything is a factor
dataAvg$subject <- as.factor(dataAvg$subject)
dataAvg$rd <- as.factor(dataAvg$rd)
dataAvg$initial <- as.factor(dataAvg$initial)
dataAvg$phonType <- as.factor(dataAvg$phonType)

# ANOVA requires a perfectly balanced design. Four subjects do not have 10 trials for every
# condition (Laura, Miguel, Teddy, Tina), so they must be dropped if minN=10
if(minN==10) {
dataAvg <- subset(dataAvg, subject != 'Laura' & subject != 'Teddy' & subject != 'Miguel'
& subject!='Tina')
}

# Compute arcsine pe
dataAvg$pAsin <- asin(sqrt(dataAvg$pe))

# Run ANOVA
aovModel <- aov(pAsin ~ rd*initial*phonType + Error(subject/(initial*phonType)), data=dataAvg)
fixedEffectsANOVA <- summary(aovModel)

# Consonants only
aovModel <- aov(pAsin ~ rd*initial + Error(subject/initial),
data=subset(dataAvg,as.numeric(as.character(phonType))>0))
fixedEffectsANOVAcons <- summary(aovModel)
```

```
# Vowels only
aovModel <- aov(pAsin ~ rd*initial + Error(subject/initial),
data=subset(dataAvg,as.numeric(as.character(phonType))<0))
fixedEffectsANOVAvows <- summary(aovModel)


############################
# Print stats
############################

# Display stats
randomEffects
fixedEffects
fixedEffectsANOVA

# Plot figure and save to PDF
sortedErrorFigure
ggsave('sortedError_consInitial.pdf')
```

Bibliography

Allen, J. B. (**2005**), *Articulation and Intelligibility* (Morgan and Claypool Publishers).

ASHA, L. (**2015**), "Brain Test Could Predict Future Literacy Challenges." ASHA Leader **20**(10), 14.

Blomert, L., Mitterer, H., and Paffen, C. (**2004**), "In Search of the Auditory, Phonetic, and/or Phonological Problems in Dyslexia: Context Effects in Speech Perception." Journal of Speech, Language and Hearing Research **47**(5), 1030 – 1047.

Bradley, L. and Bryant, P. E. (**1978**), "Difficulties in auditory organisation as a possible cause of reading backwardness," Nature **271**(5647), 746–747, URL `www.scopus.com`.

Bradley, L. and Bryant, P. E. (**1983**), "Categorizing sounds and learning to read - A causal connection," Nature **301**(5899), 419–421, URL `www.scopus.com`.

Brandt, J. and Rosen, J. (**1980**), "Auditory phonemic perception in dyslexia: Categorical identification and discrimination of stop consonants," Brain and Language **9**(2), 324–337, cited By 60.

Bronkhorst, A., Bosman, A., and Smoorenburg, G. (**1993**), "A model for context effects in speech recognition," Journal of the Acoustical Society of America **93**(1), 499–509.

Carson, K. L., Gillon, G. T., Boustead, T. M., Nippold, M., and Troia, G. (**2013**), "Classroom Phonological Awareness Instruction and Literacy Outcomes in the First Year of School." Language, Speech & Hearing Services in Schools **44**(2), 147 – 160.

Chall, J. S. (**1983**), *Stages of Reading Development* (N.Y.: McGraw-Hill Book Company).

Chenausky, K. (**2015**), "Speech in Autism Spectrum Disorder," Acoustics Today **11**.

Corriveau, K., Pasquini, E., and Goswami, U. (**2007**), "Basic auditory processing skills and specific language impairement: a new look at an old hypothesis," Journal of Speech, Language, and Hearing Research **50**, 647–666.

Cover, T. M. and Thomas, J. A. (**2006**), *Elements of Information Theory* (John Wiley and Sons, Inc.).

DeGroot, B. J., den Bos, K. P. V., der Meulen, B. F. V., and Minnaert, A. E. (**October 2015**), "Rapid Naming and Phonemic Awareness in Children With Reading Disabilities and/or Specific Language Impairment: Differentiating Processes?" Journal of Speech, Language and Hearing Research **58**, 1538–1548.

Fletcher, H. (**1929**), *Speech and hearing* (New York, D. Van Nostrand company, inc.).

Fousek, P., Svojanovsky, P., Grezl, F., and Hermansky, H. (**2004**), "New nonsense syllables database: Analyses and preliminary ASR experiments." in *Proceedings of the International Conference on Spoken Language Processing*, pp. 2749–2752.

Frey, B. J. and Dueck, D. (**2007**), "Clustering By Passing Messages Between Data Points," Science **315**, 972–976.

Fullgrabe, C. and Rosen, S. (**2016**), "On the (Un)importance of Working Memory in Speech-in-Noise Processing for Listeners with Normal Hearing Thresholds," Frontiers in Psychology **7**, 1268.

Goswami, U. and Bryant, P. (**1990**), *Phonological skills and learning to read* (Hove : Lawrence Erlbaum).

Hazan, V. and Adlard, A. (**1998**), "Speech Perception in Children With Specific Reading Difficulties (Dyslexia)." Quarterly Journal of Experimental Psychology: Section A **51**(1), 153 – 177.

Hazan, V., Messaoud-Galusi, S., Rosen, S., Nouwens, S., and Shakespeare, B. (**2009**), "Speech Perception Abilities of Adults with Dyslexia: Is There Any Evidence for a True Deficit?" Journal of Speech, Language, and Hearing Research **52**, 1510 – 1529.

Hazan, V., Messaoud-Galusi, S., Rosen, S., Schlauch, R., and Wright, B. (**2013**), "The Effect of Talker and Intonation Variability on Speech Perception in Noise

in Children With Dyslexia." Journal of Speech, Language and Hearing Research **56**(1), 44 – 62.

Hopfner, R. (**2014**), *Asymptotic Statistics: With a View to Stochastic Processes* (Walter de Gruyter GmbH, Berlin/Boston).

Joanisse, M. F., Manis, F. R., Keating, P., and Seidenberg, M. S. (**2000**), "Language Deficits in Dyslexic Children: Speech Perception, Phonology and Morphology," Journal of Experimental Child Psychology **77**, 30–60.

Johnson, C., Menon, A., Yoon, S., Han, W., Kubalanza, M., and Allen, J. (**March 2011**a), "Levels of Consosnant-Vowel Perceptual Confusion and Assessment Profiles in Children's Reading Disabilities." in *Poster presented at the Pre-Summit Workshop: Education for a Lifetime: Minds, Brains, and New Literacies.* (Innovation Summit, Beckman Institute and Office of the Chancellor, University of Illinois at Urbana-Champaign.).

Johnson, C., Menon, A., Yoon, S., Han, W., Kubalanza, M., and Allen, J. (**May 2011**b), "Levels of Consosnant-Vowel Perceptual Confusion and Assessment Profiles in Children's Reading Disabilities." in *Poster presented at the Speech Production Workshop.* (Focal Point grant, University of Illinois at Urbana-Champaign).

Johnson, C. J., Kubalanza, M. C., Allen, J., Scheidiger, C., Biller, M., and Buie, J. (**2015**), "Speech Perception and Reading Disabilities in Individual Children," in *Seminar presented at the American Speech-Language-Hearing Association* (American Speech-Language-Hearing Association, Denver, Colorado).

Kamhi, A. G. and Catts, H. W. (**2012**), *Language and Reading Disabilities* (Pearson).

Kuhl, P. K. (**2004**), "Early Language Acquisition: Cracking the Speech Code," Nature **5**, 831–843.

Li, F. (**2009**), "Perceptual Cues of Consonant Sounds and Impact of Sensorineural Hearing Loss on Speech Perception," Ph.D. thesis, University of Illinois at Urbana-Champaign.

Manis, F. R., McBride-Chang, C., Seidenberg, M. S., Keating, P., Doi, L. M., Munson, B., and Petersen, A. (**1997**), "Are Speech Perception Deficits Associ-

ated with Developmental Dyslexia," Journal of Experimental Child Psychology **66**, 211–235.

Mathworks (**2013**), *MATLAB version 8.2.0.701 (R2013b)*, The Mathworks. Inc., Natick, Massachusetts.

McDonald, J. H. (**2014**), *Handbook of Biological Statistics (3rd ed.)* (Sparky House Publishing, Baltimore, Maryland).

Melby-Lervåg, M. (**2012**), "The Relative Predictive Contribution and Causal Role of Phoneme Awareness, Rhyme Awareness and Verbal Short-Term Memory in Reading Skills: A Review," Scandinavian Journal of Educational Research **56**(4), 363–380.

Merzenich, M. M. and Jenkins, W. M. (**1996**), "Temporal processing deficits of language-learning impaired children ameliorated by training." Science **271**(5245), 77.

Messaoud-Galusi, S., Hazan, V., and Rosen, S. (**2011**), "Investigating Speech Perception in Children with Dyslexia: Is There Evidence of a Consistent Deficit in Individuals?." Journal of Speech, Language, and Hearing Research **54**(6), 1682 – 1701.

Oakhill, J. and Beard, R. (Eds.) (**1999**), *Reading Development and the Teaching of Reading* (Blackwell Publishers).

Pace, L. (**2012**), *Beginning R: An Introduction to Statistical Programming* (Apress).

Paul, P. V. and Whitelaw, G. M. (**2011**), *Hearing And Deafness: An Introduction For Health And Education Professionals* (Sudbury, MA : Jones And Bartlett Publishers).

Reetz, H. and Jongman, A. (Eds.) (**2008**), *Phonetics: Transcription, Production, Acoustics, and Perception* (Wiley-Blackwell).

Rosen, S. (**2003**), "Auditory Processing in Dyslexia and specific language impairment: is there a deficit? What is its nature?" Journal of Phonetics , 509–527.

Schuele, C. M. and Boudreau, D. (**2008**), "Phonological Awareness Intervention: Beyond the Basics," Language, Speech, And Hearing Services in Schools **39**, 3–20.

Share, D. L. (**1995**), "Phonological recoding and self-teaching: sine qua non of reading acquisition," Cognition **55**(2), 151 – 218.

Share, D. L. (**1999**), "Phonological recording and orthographic learning: A direct test of the self-teaching hypothesis." Journal of Experimental Child Psychology **72**(2), 95.

Share, D. L. and Stanovich, K. E. (**1995**), "Cognitive processes in early reading development: Accommodating individual differences into a model of acquisition," Issues in education **1**(1), 1–57.

Shriberg, L. D. and Kent, R. D. (**1982**), *Clinical Phonetics* (New York: Macmillan).

Studebaker, G. A. (**1985**), "A "Rationalized" Arcsine Transform," Journal of Speech and Hearing Research **28**, 455–462.

White, S., Milne, E., Rosen, S., Hansen, P., Swettenham, J., Frith, U., and Ramus, F. (**2006**), "The Role of sensorimotor impairments in dyslexia: a multiple case study of dyslexic children," Developmental Science , 237–269.

Ziegler, J. C., Pech-Georgel, C., George, F., and Lorenzi, C. (**2009**), "Speech-perception-in-noise deficits in dyslexia," Developmental Science , 732–745.