What /r/ Sounds Like in Kansai Japanese:
A Phonetic Investigation of Liquid Variation in Unscripted Discourse

by

Thomas Judd Magnuson
B.A., University of British Columbia, 1998

A Thesis Submitted in Partial Fulfillment of the
Requirements for the Degree of

MASTER OF ARTS

In the Department of Linguistics

What /r/ Sounds Like in Kansai Japanese:
A Phonetic Investigation of Liquid Variation in Unscripted Discourse

by

Thomas Judd Magnuson
B.A., University of British Columbia, 1998

SUPERVISORY COMMITTEE

Dr. Hua Lin, Supervisor
(Department of Linguistics)

Dr. John H. Esling, Departmental Member
(Department of Linguistics)

Dr. Tae-Jin Yoon, Departmental Member
(Department of Linguistics)

Supervisory Committee

Dr. Hua Lin, Supervisor
(Department of Linguistics)

Dr. John H. Esling, Departmental Member
(Department of Linguistics)

Dr. Tae-Jin Yoon, Departmental Member
(Department of Linguistics)

ABSTRACT

Unlike Canadian English which has two liquid consonant phonemes, /ɹ, l/ (as in *right* and *light*), Japanese is said to have a single liquid phoneme whose realization varies widely both among speakers and within the speech of individuals. Although variants of the /r/ sound in Japanese have been described as flaps, laterals, and weak plosives, research that has sought to quantitatively describe this phonetic variation has not yet been carried out. The aim of this thesis is to provide such quantification based on 1,535 instances of /r/ spoken by four individuals whose near-natural, unscripted conversations had been recorded as part of a larger corpus of unscripted Japanese maintained by Dr. Nick Campbell of Advanced Telecommunications Research Institute International (ATR), Kyoto, Japan.

Tokens of /r/ were extracted from 30-minute conversations between one pair of male speakers and one pair of female speakers. Each token was narrowly transcribed into the International Phonetic Alphabet, then categorized based on the author's perception of: 1) the strength/narrowness of central oral articulatory stricture, and 2) the presence or absence of an auditory-perceptual lateral and/or rhotic sound quality. Transcription and category frequencies for each speaker averaged across all environments were then compared with frequencies in specific phonological environments to ascertain whether a particular environment was amenable to a 'drift' towards any particular category of variant, and whether patterns of 'drift' applied to all speakers or varied on an individual basis. Transcriptions of the 1,535 tokens of /r/ ranged widely among lateral and non-lateral flaps, raised (i.e. increased articulatory contact) non-lateral flaps akin to light voiced plosives (e.g. Hattori 1951, Kawakami 1977), as well as lateral approximants and rhotic approximants. While two of the four speakers, both males, patterned similarly by dividing their productions of /r/ chiefly among short lateral approximants and rhotic approximants, each speaker did vary considerably in their choice of variants in any given environment. Drift is considered in terms of physiological parameters which may be optionally exploited to maintain phonological salience.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ACKNOWLEDGMENTS

Were it not for the knowledge, advice, support, and wisdom of a great many people, the work presented here would not have been possible. First among these is Dr. Hua Lin, my thesis supervisor. Dr. Lin's approach to guiding her students to the completion of their projects is exceptional in that she is able to allow them to fully experience the process of learning. Learning involves making mistakes and taking detours which, even though they may not be written into the ink of a thesis, are nonetheless critical vectors for the transmission of knowledge that cannot necessarily fit between lines of triple-spaced print. My thanks go to Dr. Lin for allowing me the latitude to learn.

I also wish to thank Dr. Nick Campbell and his colleagues for granting me access to the invaluable speech data recordings on which this study is based. Equally important to this study has been the advice and guidance provided by my Committee members: Dr. John Esling and Dr. Tae-Jin Yoon, whose combined technical and practical assistance helped me program a computer while not forgetting to "mind the pointy bits." Finally, many thanks to UVic Linguistics graduate students, faculty, and staff including Dr. Leslie Saxon, Soo-Youn Ham, Scott Moisik, and many others whose support and advice has been invaluable.

For Mom & Dad

Chapter One

INTRODUCTION

1.1 Study motivation

The author's motivation for this study comes in part from reflections on several years' language teaching, translation, and life experience spent in various regions of Japan. While it is well-known that many Japanese-speaking learners of English as a foreign language have difficulty with the *l*- and *r*-sounds of English (e.g. Shimizu & Dantsuji 1987, among others), it is perhaps less well-known that this difficulty is not unidirectional. Native speakers of North American English (NAE) learning Japanese as a foreign language, including this author, often require much time to master the *r*-sound in Japanese (e.g. Kokken 1990, Akamatsu 1997). These cross-linguistic difficulties in mastering /r/ are, moreover, echoed in first language acquisition. Similar to how North American English-learning infants stumble between r's and w's in their acquisition of their *r*-sound, /ɹ/ (e.g. McGowan et al. 2006), infants learning Japanese often confuse their *r*-sound, the alveolar or post-alveolar flap /ɾ/ (as in NAE *ladder*, or *water*), with /d/ (e.g. Otsuka 1991). What is striking about this is not so much that these difficulties in acquisition exist, but more that the variety of speech sounds that can be referred to as '*r*-sounds' is extensive to the point that 'r' can be confused with sounds as different as /d/, /l/, and /w/. One of the motivations for this study stems from the author's interest in exploring, as a learner of Japanese, the possible diversity of speech sounds grouped under the label /r/ in Japanese.

1.2 Why is sub-phonemic variation in Japanese flaps worth looking at?

It is often pointed out that, in any number of fields from business and economics to anthropology and human ecology, improved communications technologies and the ability to interact in real-time with people from diverse language and cultural backgrounds has made the world a much smaller place. For linguistics, this shrinking of the global community has triggered interest in developing ways for people to more easily talk to one another. Not only internationally but within the contexts of multi-cultural and multi-lingual societies as well, entire sub-disciplines such as applied linguistics, foreign language acquisition, as well as automated speech recognition and synthesis have taken up the task of developing both models to explain and means to better facilitate speech communication. These models and methods, in turn, draw from a long and evolving history of research into how the sounds of speech themselves are related to one another in terms of how they are produced (their phonetic properties) and how they pattern in language (their phonological properties). Until recently, however, much of the research available to teachers, language facilitators, and speech technologists has been based largely on speech data compiled under controlled laboratory circumstances. The data that have traditionally fed our models and methods, that is, has not been natural speech but controlled speech - idealized versions of how speech is actually realized in daily human interaction. While the investigation of controlled speech is extremely valuable and necessary for defining what acoustic and/or articulatory *targets* are involved in producing the phonemes of a language, the extent to which those targets are deviated from in natural speech, and the possible reasons for that deviation, ought not be overlooked. That is, understanding how speech sounds vary can give us further insights into how speech sounds are organically related to one another in terms of their physical production as well as how they are affected by the context (linguistic, pragmatic, social, and otherwise) in which they are spoken. This study focuses on near-natural, unscripted speech in order to add to the body of data from which language researchers and developers can draw, so that the range of variation among speech sounds in nearer-to-natural speech can be better incorporated into language theory and practice.

There are a number of areas where making data on variation in near-natural speech available to the research community can be potentially beneficial. One of these areas is automated speech recognition and synthesis. Anyone who has ever telephoned a large enterprise has likely interacted with 'robot operators,' which are not always adept at parsing our requests. One of the reasons that these systems do not always perform to our expectations is that, arguably, they may be programmed to recognize a single speech sound according to a range of variation that is limited in scope to what is phonologically plausible. For instance, a speech recognition system designed with NAE in mind might be trained to recognize the brief interruption in an acoustic signal made by a flap [ɾ] as an instance of /t/ or /d/, as flaps are predictable allophones. However, a speaker of Japanese who is trying to communicate in English with the same speech recognition device might produce a flap intending on /r/. In this case, if the system is not primed to recognize flaps as possible variants of /r/, there could be misunderstanding. What's more, assuming that natural speech entails a greater degree of variation than careful speech or citation forms, it is important first to document that variation so that speech recognition systems can be trained to accept those variants and map them on to speakers' intended targets. With Japanese /r/ in particular, which has been described as being subject to variation among rhotic approximants, flaps, and laterals (e.g. Hattori 1951, Kawakami 1977, Vance 1987, Akamatsu 1997, Okada 1999), enumerating the variants is particularly valuable. As this enumeration has not yet been documented, the primary goal of the present study is to provide qualitative observations about /r/ in Japanese along with quantitative descriptions of frequency, which have not been featured in previous research. The main task, that is to say, is the quantitative documentation of qualitative variation.

1.3 Research questions

While the literature informing this study's research questions will be explored in detail in Chapter 2, the questions themselves are listed in Table 1 below followed by a brief introduction to their rationale.

Table 1. Research questions

Q. 1:	What speech sounds occur for Kansai Japanese /r/ in extemporaneous conversation?
Q. 2:	Do lateral variants occur more frequently when adjacent to the vowels /a, e, o/?
Q. 3:	What variants of Kansai Japanese /r/ occur post-pausally?
Q. 4:	What variants of Kansai Japanese /r/ occur following the nasal /n/?
Q. 5:	To what extent do speakers pattern similarly in their choice of /r/ variants?
Q. 6:	How do the observed variants of /r/ in Kansai Japanese relate to our understanding of rhotics in general?

The first research question is perhaps the most central to the motivation of this study: what sorts of sounds, that is, can be observed in the near-natural speech available to the study? The literature (e.g. Hattori 1951, Kawakami 1977, Vance 1987, Akamatsu 1997, Okada 1999) suggests a range of sounds from weak voiced plosives, lateral and rhotic approximants, taps, as well as palatalized realizations of the foregoing. Some of these authors suggest that the vowels /a, e, o/ are particularly amenable to laterals (Q. 2), and the same authors agree that post-pausal, or word-initial /r/ is realized as a weak plosive (Q. 3). Following nasals (Q. 4), some of the literature points to weak plosives while other authors suggest that laterals are common. Question 5, asks about the degree to which variation in speakers' /r/s is individual; that is, the one other point of consensus in the literature reviewed in Chapter 2 is that any variant may be observed, "depending on the speaker and the situation" (Akamatsu 1999:106). Lastly, Q. 6 discusses how the /r/ variants observed in the data fit into our understanding of rhotic speech sounds in general, as exemplified by the presence of multiple parameter associations as per the *family resemblances* model proposed by Lindau (1985).

1.4 Organization of this thesis

This thesis is organized as follows. Chapter 2 provides an introduction to the sounds of Japanese as well as an overview of the language as it is spoken in the Kansai region. Also, literature that discusses rhotic speech sounds, and Japanese /r/ in particular, is discussed in further detail. Chapter 3 introduces the methodology used in this study, and its results are given in Chapter 4. This thesis concludes with a summary and discussion of the results in Chapter 5.

Chapter Two

LITERATURE REVIEW

2.1 The sounds of Japanese

Table 2 gives a listing of the consonant inventory of Japanese, as adapted from multiple sources including Vance (1987), Shibatani (1990), Akamatsu (1999), Okada (1999), and Grenon (2005).

Table 2. Consonant sounds in Standard Japanese (Vance 1987; Shibatani 1990; Akamatsu 1999; Okada 1999; Grenon 2005). Brackets indicate allophones; where two symbols appear together, the sound to the left is voiceless and that to the right is voiced.

	Bilabial	Alveolar	Post-alveolar	Palatal	Velar	Glottal
Stops	p b	t d			k g	(ʔ)
Fricatives	(ɸ)	s z	(ʃ)	(ç)		h
Affricates		(ʦ) (ʣ)	(ʧ) (ʤ)			
Approximants	w			j		
Liquids		r				
Nasals	m	n		(ɲ)	(ŋ)	

The consonants which appear in brackets represent allophonic variants, while those without brackets represent the phoneme inventory as suggested in Shibatani (1990:159). The consonants in Table 2 are complemented by five vowel phonemes: /i, e, a, ɯ, o/, of which /i, ɯ/ are subject to devoicing between voiceless obstruents (Kubozono 1999, Shibatani 1990:161). Shibatani (1990) additionally lists a homorganic nasal /N/ and moraic obstruent /Q/ as phonemes for Japanese. Phonological accounts of Japanese (e.g. Shibatani

1990, Kubozono 1999, Ito & Mester 1999) generally view the moraic obstruent /Q/ as taking on the place and manner features of following syllable-initial obstruents as in the word *hakkiri* ('clear'), optionally phonemicized as /haQkiri/. The nasal /N/, for its part, similarly assimilates to the place of following obstruents. For the purposes of this thesis, and insofar as the effect of a putative moraic obstruent is gemination, it is left out of the inventory here in favour of the lengthening diacritic [ː] which will be used to express length contrasts hereafter. As for the nasal, and although there may indeed be phonological grounds for claiming phonemic status of a homorganic /N/, /n/ is assumed to be the phoneme here for the sake of simplicity; positional variants (or what would otherwise be taken as assimilatory variants of /N/) are listed alongside /n/ in Table 2.

Although adapted from literature on standard Japanese, the inventory in Table 2 is arguably also applicable to Kansai Japanese with the exception of the high back unrounded vowel /ɯ/, which Shibatani (1990) suggests is slightly more rounded in Kansai dialects than it is in Standard (Tokyo) Japanese. Stemming from phonological processes such as palatalization and nasal assimilation, both of which occur widely in Japanese, there is much debate over the phonemic status of many of the consonants listed in Table 2. For instance, as to the status of [z] and [ʥ], Grenon (2005) observes that both sounds occur phonetically in the language, but argues for the phonemic status of /ʥ/ over [z], which she claims is an allophone based on patterns of lenition. The variety of speech sounds that are possible in the liquids category highlighted in Table 2 is the focus of the present study, and although trills (IPA symbol [r]) can and do occur in the language, the claim here is not that [r] is phonemic. Rather, the symbol 'r' in the table is here used as a label of convenience borrowed from the conventional representation of the sound in Romanized orthographies of Japanese. While this phoneme is most commonly thought of as an alveolar or post-alveolar tap (IPA [ɾ]) (e.g. Hattori 1951, Kawakami 1977), Akamatsu (1997) argues for an additional palatalized tap ([ɾ'], IPA: [$ɾ^j$]) that occurs phonemically in Sino-Japanese lexical items (see Ito & Mester 1999 for discussion) such as *shoryaku* ('abbreviation') and *ryokou* ('travel').

2.2 Standard versus Kansai Japanese

What follows is a brief overview of the Japanese language and the dialect, or more properly group of dialects, referred to as 'Kansai Japanese.' The notion of a 'standard' dialect of Japanese is a relatively new one in historical terms, stemming from the nation-building and modernization processes which began in earnest following the Meiji Restoration in 1868. This event marks the division of Japanese history between the Early Modern Period and the Modern Period (Beasley 1972:1-3). The 'restoration' which occurred was the ostensible return of political power to the Emperor and end to the rule of the Tokugawa Shogunate, which had ruled the country as a series of strictly separate feudal domains among which travel and communication were severely restricted by the central government (the Shogunate, or *Bakufu*). The Bakufu had governed the country from Edo (present-day Tokyo, see Figure 1); however, it drew its political legitimacy from the

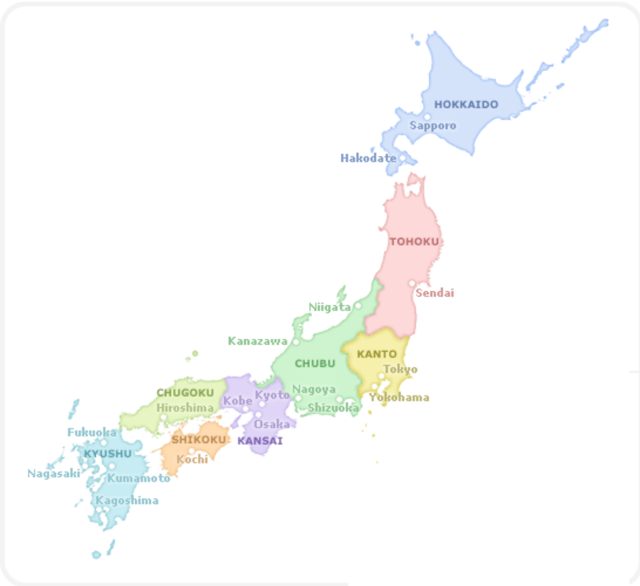


Figure 1. Regional map of Japan (JNTO 2007, reprinted with permission)

Emperor whose court had been located in Kyoto since the Heian Period. Following the Meiji Restoration, the imperial court was moved from Kyoto to the renamed Tokyo, and the theretofore segregated feudal domains were re-organized into a newly formed national polity as prefectures. The linguistic legacy of feudal domains, however, remained: each domain had its own dialect(s), the differences between which having been enhanced through their mandated isolation. Citing an example from Hattori (1960), Gottlieb (2005:39) points out that the forms of Japanese spoken in Kagoshima (southern Kyushu) and Sendai (eastern Honshu) were mutually unintelligible.

Given the post-Meiji Restoration need to unify the former domains into a single national state, the development of a variety of Japanese which could be understood in all corners of the archipelago gained currency among intellectuals, writers, and political leaders in the decades before 1900. As of 1901, the Japanese taught to school children throughout the country was designated by the Ministry of Education “to be that of middle- and upper-class Tokyo residents;” moreover, in 1916, ‘Standard’ Japanese was defined by the National Language Research Council (a.k.a *Kokken*, est. 1902) as the speech of educated residents of the Yamanote district of Tokyo (Gottlieb 2005:7-8). In subsequent decades, Standard Japanese would continue to be taught through the public education system (forcibly so for much of the 20th century), and would be further broadcast and prescribed by the national broadcaster, NHK (Gottlieb 2005:9).

Similar in some ways to Standard Japanese, Kansai Japanese is also a relatively recent invention. The term *Kansai* is both a geographical and psycho-social marker which, set in opposition to *Kanto*, delimits western and eastern centres of culture and industry. While Tokyo and its environs can be referred to collectively as Kanto, Kansai can be used to talk about Osaka and its surrounding cities. From a purely geographical perspective, the region that Kansai generally maps onto is called Kinki (see Figure 2), and includes (from west to east) the major urban centres of Kobe, Osaka, Kyoto, and Nara.

Figure 2. Map of the Kinki/Kansai region (Shingu & Hatanaka 2007, reprinted with permission)

In part owing to the historical, social, economic, and political importance of the region, the speech of this part of Japan enjoys a high degree of covert prestige (Labov 1966) insofar as it remains a desirable marker of regional identity. Richmond (2004:70), for instance, identifies the Kyoto dialect as being widely perceived (or rather stereotyped) as graceful and polite through an amalgam of socio-culture attitudes held by Japanese speakers in general. The neighbouring dialect of Osaka enjoys a similar positive stereotype, though for different reasons. As opposed to politeness and grace, Osaka Japanese is arguably viewed as being direct and colloquial - very much in the same vein as the working class speech of New York City residents studied by Labov (1966). Insofar as the dialects of Kyoto and Osaka both are considered Kansai dialects, it can be argued that its covert prestige stems from both high and low culture: the former cultural and political power of Kyoto as well as the 'salt of the earth' directness of Osaka. That is, although the political capital of the country had been in Edo, and then Tokyo as described above, the Kansai region had been the seat of Imperial authority for many centuries: Kyoto until 1868, and before that Nara. In modern times, the Osaka dialects of popular entertainers such as Akashi Sanma and the comedy duo Downtown (a.k.a. Hitoshi Matsumoto and Masatoshi Hamada) have found national television, radio, and cinematic audiences. The covert prestige of Kansai Japanese, then, is fed by a collusion of the allure of traditional Kyoto culture and the visibility of Osaka popular culture. Recent years have also seen the growth of Kansai-Japanese teaching materials designed for both learners of Japanese as a foreign language (e.g. Shingu & Hatanaka 2007, Palter & Slotsve 2006) and native Japanese speakers of other dialects (e.g. Kawauchi 1993, Okamoto et al. 1998).

2.3 Rhotics as a class of speech sounds

Some types of speech sounds such as stops (e.g. [p, t, k]), fricatives (e.g. [f, s, x]), and vowels (e.g. [a, i, o]) can be grouped together based on the fact that they share common articulatory or acoustic properties. For instance, articulatorily speaking, stops involve a complete blockage of airflow in the vocal tract while vowels can be defined by their lack of blockage of the same airflow save for some configuration of the supra-glottal passages which does not result in audible friction (e.g. Sweet 1906). Fricatives, for their part, involve the formation of a stricture (or blockage) in the vocal tract such that the air passing through the stricture becomes turbulent, resulting in what is perceived by a hearer as frication *noise* (e.g. Catford 1977). In terms of their acoustic properties, the open state of the vocal tract for vowels allows the periodic vibration of the vocal folds to resonate through the vocal tract and be shaped, and not blocked, by different configurations of the throat and tongue. This gives vowels their characteristic, sonorous resonant frequencies, or *formants*. Stops and fricatives block, or obstruct these frequencies while other types of speech sounds such as glides and approximants momentarily warp these frequencies as the tongue 'glides' from one open configuration to another. In these ways, insofar as the articulatory and/or acoustic properties of speech sounds form relations amongst themselves independent of the languages that might deploy them to convey meaning, phonologists have used these relations to divide speech sounds into *natural classes* (e.g. Halle 1961; Odden 2005).

The collection of speech sounds referred to by the labels *rhotics*, or *r-sounds*, is more troublesome in that no single articulatory or acoustic property has been identified which binds them together as a class of sounds (e.g. Lindau 1985, Ladefoged & Maddieson 1996). That is, the term *rhotic* has been used to classify a wide diversity of speech sounds with different manners and places of articulation, as well as acoustic properties. The rhotics examined in Lindau's (1985) study (discussed in further below), for instance, ranged in place of articulation from alveolars and post-alveolars to velars and uvulars. In terms of manner, these ranged from approximants and fricatives to trills and taps. Rhotics' cohesion as a group of speech sounds has in fact been the product of orthographic convention and phonological observation. Ladefoged & Maddieson (1996:214) note that part of why these sounds find themselves together is that, historically in orthographies flowing from the Greco-Roman tradition, they have been symbolized in writing using various derivatives of the letter 'r.' The representation of these various sounds by the symbol 'r' has led, as Walsh Dickey (1997:13) points out, "to work in pedagogy and philology which does not question the diversity of sounds that fall under the rhotic rubric."

The treatment of rhotics in linguistic analyses has been to place them within the class of *sonorants* along with sonorous laterals, which together form the sub-category of *liquids* (Walsh Dickey 1997:16). The motivation for grouping rhotics and laterals together is not ad-hoc; phonologically rhotics and liquids behave in similar ways and often occur in environments where other natural classes of sounds do not, such as adjacent to obstruents in consonant clusters. The *r*-sound in the word *obstruent* and the *l*-sound in the word *cluster* are examples of such a 'phonologically privileged' position. There is, moreover, at least one speech sound that combines the phonetic characteristics of both laterals and rhotics. The voiced alveolar lateral tap [ɺ], which Walsh Dickey (1997) argues patterns as a subset of rhotics, occurs in only 3.4% of the world's languages (UPSID 1992). Ladefoged & Maddieson (1996:243) suggest that, phonetically, lateral taps are produced when one side of the tongue remains down during the gesture for a tap, thereby opening a lateral channel for airflow to escape the vocal tract. They also suggest that, in languages where [ɾ, l] alternate like Nasioi, Barasano, and Tucano, back vowels are generally amenable to the production/perception of laterals while front vowels prefer rhotics.

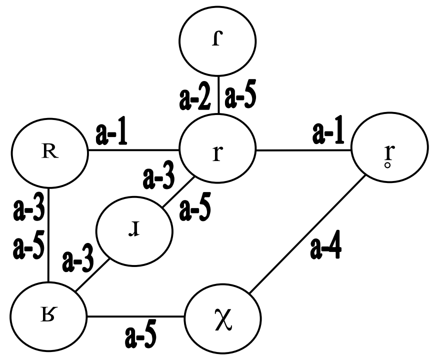
2.3.1 The family resemblances model (Lindau 1985)

The search for a single phonetic parameter which can be claimed as unifying rhotics as a class of sounds motivated Lindau's (1985) detailed study of *r*-sounds from a wide variety of languages. One of the candidates for being such a parameter was the lowering of the third formant frequency (F3) - a characteristic that has been widely associated with the rhotic approximant [ɹ] in North American English (e.g. Delattre 1968; Westbury et al. 1998; Espy-Wilson et al. 2000). Lindau's investigation looked at some 9 different /r/s in 15 languages and dialects spoken by over 90 speakers. What she found was that although a lowered F3 was characteristic of [ɹ] in North American English, it was quite rare cross-linguistically. Uvular *r*-sounds such as the uvular trill [ʀ][1] in French and uvular fricative [ʁ] in Southern Swedish, for instance, had high F3 values. The conclusion that Lindau draws is that there is in fact no single articulatory or acoustic parameter that binds rhotics together as a class; rather, they are associated with one another through a web of *parameter relations* or *family resemblances* (Lindau 1985:165-7; she attributes the term 'family resemblances' to Wittgenstein 1958).

Lindau's parameter relations model of rhotics is shown in Figure 3. Lindau posits five parameters to link rhotics together: *pulse pattern* (a-1) describes trills' ([ʀ, r, r̥])

[1] Lindau uses the symbol [ř].

characteristic repetition of blockages (or ‘pulses’) of the airflow through the vocal tract. *Closure duration* (a-2) refers to the brevity of each pulse in a trill and analogizes this to the brief closure formed in taps/flaps [ɾ]. The *presence of formants* (a-3) indicates that there are characteristic quasi-periodic frequencies as for vowels and other sonorants; with trills, these occur between pulses. *Presence of noise* indicates frication noise, as for the voiceless uvular fricative [χ] and devoiced trill [r̥]. Lindau’s last parameter is *spectral energy distribution*, an acoustical representation of place of articulation.



a-1	Pulse pattern (trill)
a-2	Closure duration
a-3	Presence of formants (sonorant)
a-4	Presence of noise
a-5	Spectral energy distribution (place)

Figure 3. Family resemblances/parameter relations among rhotics (Lindau 1985:167)

The thrust of Lindau’s model is that, while there is no single parameter that is common to all rhotics, all rhotics do have some parameter(s) in common with others, which in turn are linked to yet other rhotics by virtue of further parameters. For instance, taps and the rhotic approximant [ɹ] seemingly do not have anything in common; however both sounds have different features in common with trills (i.e. trills share the presence of formants with approximants and closure duration with taps). In this way, rhotics as a group of sounds are akin to a quilt made from threads of different colours and thicknesses: one thread interweaves with another, which then interweaves with another.

While Lindau's model succinctly interprets rhotic parameter associations in acoustic terms, it can be critiqued insofar as it does not fully incorporate features of speech production that originate lower in the vocal tract than the oral cavity. Earlier work by this author proposed a recasting of rhotic parameter relations (Magnuson 2007b, Figure 4) in order to incorporate the functioning of the laryngeal-pharyngeal vocal tract (LPVT) to complement the family resemblances model. Specifically, an additional articulatory parameter, *pharyngeal modification*, was suggested alongside *aryepiglottic trilling* and *vocal fold vibration*. A schematic depiction of Magnuson (2007b) is given overleaf in Figure 4.

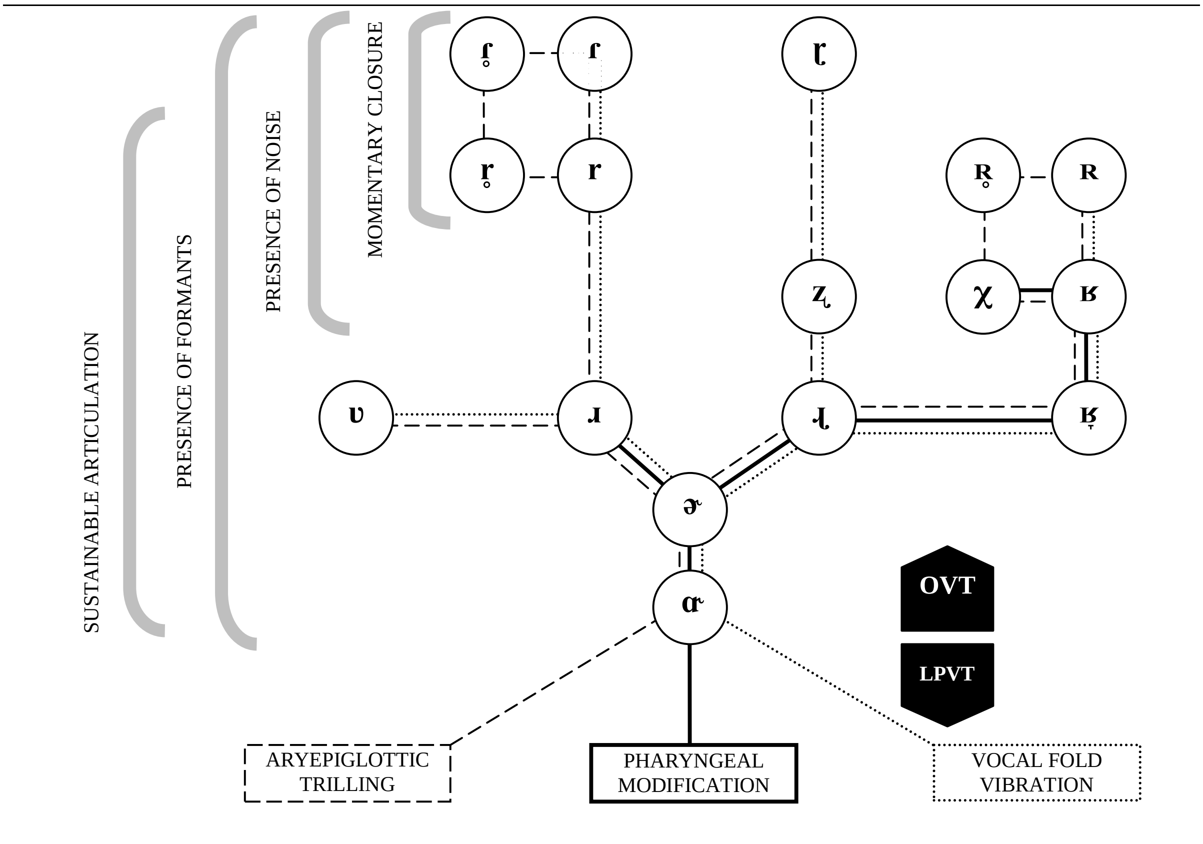


Figure 4. A recasting of rhotic parameter relations to incorporate the laryngeal-pharyngeal vocal tract (Magnuson 2007b). OVT= oral vocal tract; LPVT= laryngeal-pharyngeal vocal tract

Aside from the three LPVT parameters, all of which are proposed as being gradient, the model in Figure 4 reinterprets Lindau's *pulse pattern* and *closure duration* parameters in terms of articulatory sustainability. Based on Edmondson & Esling's (2006) valve model of the LPVT, the parameter *pharyngeal modification* is conceptualized as potentially involving any combination of tongue-root retraction and/or pharyngeal narrowing (Valves 4 and 6 in Edmondson & Esling's model). This parameter arguably associates rhotics, such as

American English /ɹ/ (e.g. Delattre & Freeman 1968; Espy-Wilson et al. 2000) which involve some degree of pharyngeal constriction alongside a primary OVT articulation.

2.4 What is a [ɾ]*?*

Articulatorily speaking, the speech sound corresponding to the IPA symbol [ɾ] is a rapid gesture involving the front of the tongue and the alveolar region of the mouth. Following from an articulatory distinction identified by Ladefoged (1964) and echoed by Catford (1977), among others, two names have traditionally used in reference to this symbol: *tap* and *flap*. *Taps* involve a ballistic gesture wherein the tip of the tongue momentarily comes into contact with (i.e. taps) the region of hard tissue behind the top front teeth, the alveolar ridge, then immediately returns to its former position. *Flaps*, in contrast, involve the front part of the tongue coming into contact with the alveolar ridge *on its way* to a new position consistent with whatever speech sound is to follow (Catford 1977).

Recent ultrasound research by Derrick & Gick (2008) suggests that, at least in North American English where [ɾ] is an allophone of post-stress /t, d/, there may in fact be a four-way sub-phonemic but categorical distinction in the way [ɾ] is articulated (Figure 5). They observe *up-flaps* and *down-flaps*, as well as *alveolar taps* and *post-alveolar taps* depending on what combination of vowels and vocalic 'r' ([ɝ]) constitute the tap/flap's environment.

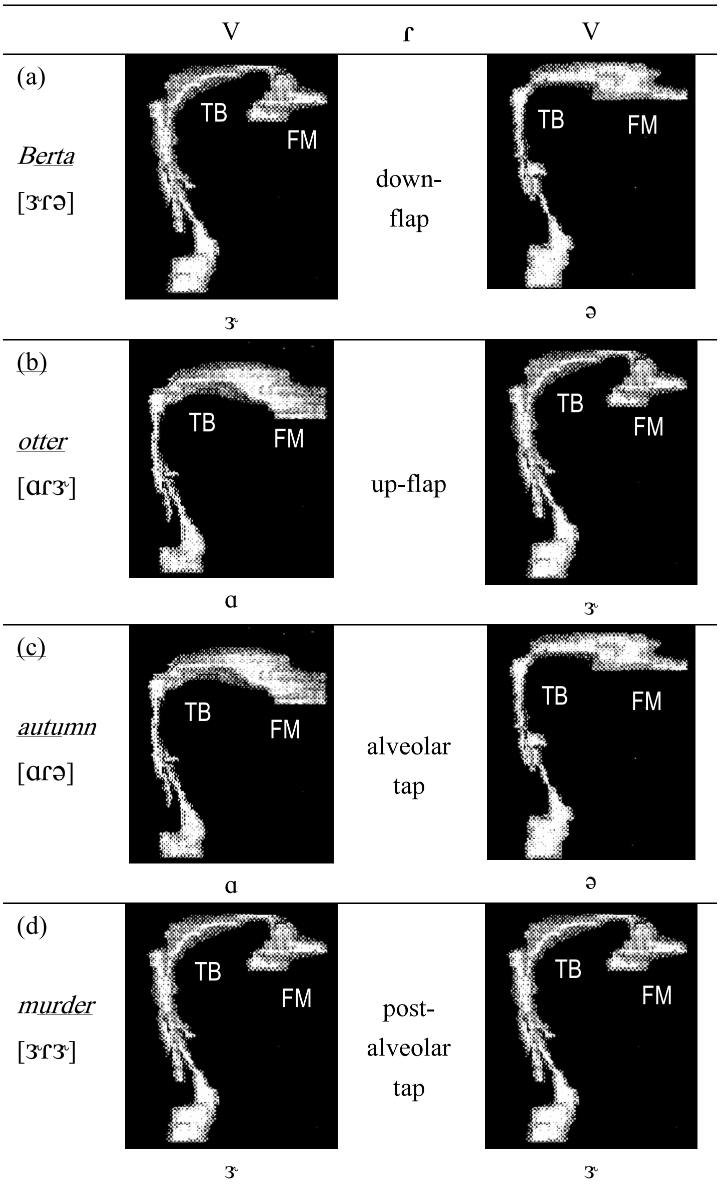


Figure 5. Four categories of flaps/taps in NAE (Derrick & Gick 2008; vocal tract MRI images from Story et al. 1996, adapted). TB = tongue body; FM = front of mouth

Specifically, they claim that since vocalic ‘r’ involves a tongue position such that the tongue tip is above the alveolar ridge and other vowels involve a tongue tip position below the alveolar ridge, the kinematic gesture between a vowel and vocalic ‘r’ will be a resolution of that positional conflict. *Down-flaps* (Figure 5(a)) occur where vocalic ‘r’ precedes a flap-vowel sequence, as in the name *Berta*; the tongue tip contacts the avleolar ridge from above, then moves into a position consistent with the following vowel, a schwa. *Up-flaps,* as in the word *otter* (Figure 5(b)) involve the inverse pattern: the tongue tip strikes the alveolar ridge from below on its way to assume a position for vocalic 'r' above the alveolar ridge. Derrick & Gick say that *alveolar taps* occur between vowels, as in the word *autumn* (Figure 5(c)), where the tongue tip touches the alveolar ridge then immediately returns to its former position. *Post-alveolar taps* occur in words such as *murder* (Figure 5(d)), where the consonant falls between two vocalic 'r's. Since the tongue tip is situated above the alveolar ridge prior to and following the consonant, contact is made by the tongue-tip and the post-alveolar region. While these kinematic distinctions between flaps and taps are returned to later in this thesis, where the distinction is not salient to the discussion the term *flap* will be used hereafter to refer generally to all speech sounds encompassed by the symbol [ɾ].

Phonologically, flaps as speech sounds can be thought of as jacks-of-all-trades. In North American English, as discussed above, flaps can take the place of alveolar stops in words such as *little*, *water*, and *ladder* – where the syllable before /t, d/ is stressed and the following one is unstressed (e.g. Zue & Laferriere 1979, Patterson 2001, Fukaya & Byrd

2005). While flaps can function as stops as in the case of North American English and, less famously, in Xiangxiang Chinese (Zeng 2007), the perhaps more common role of flaps is as rhotic liquids, or /r/s. Japanese, the language investigated here, is a case in point; however, other languages with flap /r/s include (among many others) Hausa (e.g. Ladefoged 1964), Spanish (e.g. Monnot & Freeman 1972), Catalan (Recasens & Pallarѐs 1999), Korean (e.g. Cho 1969, Joo 2005), and Tamil (e.g. McDonough & Johnson 1997, Narayanan et al. 1999). Among languages that have rhotic flaps, there is often alternation with other kinds of rhotics as well; they have been observed to commonly alternate will trills ([r]) as in Spanish and Catalan, and laterals ([l]) as in Korean. While flaps have been discussed as reduced variants of alveolar stops in North American English, work done by Natasha Warner and colleagues with near-natural, extemporaneous speech shows that these reduced forms are often further reduced to approximants as well (e.g. Tucker & Warner 2007). This author's own preliminary comparison of flaps in Canadian English and Japanese (Magnuson 2007a) similarly suggested that a range of variation, from plosive-like realizations to approximants, is possible in how flaps are produced.

2.5 The behaviour of flaps in other languages

Japanese is not alone in having a flap as a rhotic. Considering flaps together with taps, Walsh Dickey (1997:15) points out that, according to the UPSID (1992) database, these make of 40.1% of rhotics in the world's languages. Although flaps/taps may be the only contrastive liquid in some languages, such as in Japanese and Korean (e.g. Cho 1967), other languages such as Catalan (Recasens & Pallarès 1999) and Tamil (e.g. McDonough & Johnson 1997, Narayanan et al. 1999) contrast flaps with multiple rhotics in the same language - including different types of flaps as in the case of the 5-way liquid contrast in Tamil. Even where flaps do not occur as phonemes in a language, they are frequently observed as allophonic variants of other speech sounds. Perhaps the most well-known case of this is the flap allophone of North American English /t, d/ where plosives are said to lenite to flaps following a stressed syllable (e.g. Zue & Laferriere 1979, Patterson 2001, Warner 2007 (c.f. Fukaya & Byrd 2005)). Among rhotic phonemes, flaps or taps often occur as allophones of trills in a variety of languages including Spanish (e.g. Monnot & Freeman 1972), as well as several varieties of Brazilian Portuguese (e.g. Harris 2005).

2.6 Some different views of /r/ in Japanese

This section will describe a number of views in the literature on how /r/ in Japanese is realized. It is worth mentioning beforehand, however, that these views represent interpretations of /r/ in Standard (Tokyo) Japanese, and that other dialects' /r/s are considered very little in the literature on the whole. That having been said, the observations of the authors presented below are invaluable in understanding /r/ variation in that they reflect decades of impressions made by phonetically trained ears on the natural speech of individuals pursuing the course of their daily lives, free from constraints imposed by any sort of experimental design. The upside of this, of course, is that the observations are based on 'real' language; the downside to this, however, is that these observations about natural speech have not yet been quantitatively verified.

A common theme among the authors' observations described below is that there is a great deal of individual and stylistic variation in how /r/ is realized. They also identify a number of phonological parameters which condition how variants pattern. Generalizing over slight differences between the authors' points of view, these parameters include: 1) post-pausal versus intervocalic context, typified by greater or less degree of stricture, 2) whether or not /r/ occurs before the vowel [i], resulting in palatalization, and 3) whether or not /r/ follows [n] or precedes [a, e, o], in which case laterals are common.

2.6.1 Amanuma et al. (2004)

Written for a Japanese-speaking readership, Amanuma et al. (2004) is an introductory book in Japanese phonetics; it is included here as it provides a glimpse of how the topic of liquid variation is introduced to an undergraduate audience. They suggest seven possible realizations for /r/: non-palatalized [ɾ, l, lː, r] and palatalized [ɾ̣, ḷ, ḷː] (current IPA: [ɾʲ, lʲ, lʲː]). With the proviso that they are casting /r/ in general terms and that there is a great variety of ways that people produce their *r*-consonants depending on, for instance, whether they feel nervous or are speaking casually, they say that flaps occur word-internally, and presumably between any two vowels other than /i/ (before which the palatalized variant [ɾʲ] would be expected). Word-initially (and presumably post-pausally) as well as following [n] within words, the lateral approximant [l] occurs. The authors do not indicate whether or not this [l] is the same as the English non-velarized [l], although both are discussed in the same section of their text. Without being specific in numerical terms, they also suggest that many speakers of Japanese exclusively use the laterals [l, lʲ] in their speech to the exclusion of other variants. The symbols with the lengthening diacritic [ː] apply to the sounds when /r/ is spoken as a geminate consonant, which does not occur normally in the lexicon though it can be prompted orthographically using the kana characters っ or ッ in lower case before another kana character which includes /r/ as its consonantal component. Finally the authors note that the trill [r] occurs very rarely, and is restricted to rough/highly casual (*beranmechou* is the word they use) speech (Amanuma et al. 2004:74-76).

2.6.2 Okamura (1995).

Okamura (1995) gives another general picture of /r/ within the context of an over-view chapter on Japanese phonology and phonetics in Kindaiti et al. (eds.) *An Encyclopaedia of the Japanese Language*. Okamura says that, generally, intervocalic /r/ is an apical tap [ɾ]. At the beginning of words, Okumura describes /r/ as being optionally realized as the lateral approximant [l] or a "light plosive" for which no symbol is given (Okamura 1995:247).

2.6.3 Kawakami (1977)

Kawakami's (1977) *Nihongo Onsei Gaisetsu* (Overview of Japanese Phonetics) is just that: an overview of Japanese phonetics, and his view of /r/ is echoed in the English literature by Vance (1987). Kawakami describes the "most representative" realization of /r/ as an apical tap [ɾ], with the lateral [l] also a possibility, especially before the vowels [a, e, o] and when /r/ is produced as a geminate. He suggests here that, even among speakers not given to produce /r/ as a lateral may do so where prompted to say an interjection such as *arre, maa* ('What's that? Well…'). Word-initially Kawakami describes /r/ as a weak plosive akin to [d]. In terms of the articulatory gestures involved in producing the tap /r/, he argues that, insofar as the sound of /r/ results from the apex of the tongue quickly leaving the gums, the motion of the tongue tip prior to making contact is inconsequential. Given this, the weak plosive found word-initially and the tap arguably produce the same auditory effect.

Though he does not identify these as separate phonemes as such, Kawakami notes palatalized '[rj]' occurs as the consonantal component before [a, ɯ, o] in the orthographic symbols りゃ, りゅ, りょ (*rya, ryu, ryo*). Insofar as he describes the otherwise non-palatalized '[r]' as occurring before the vowels [a, i, ɯ, e, o], we can say that palatalized /r/ is contrastive at least where it is found in front of the vowels [a, ɯ, o].

2.6.4 Hattori (1951)

Hattori' (1951) *Onseigaku* (Phonetics) is an often-cited reference in phonetics literature written in Japanese, and is also one of the few that discusses (albeit briefly) what /r/ variants are realized in dialects other than that of Tokyo. In the Kumamoto dialect on the southern island of Kyushu, for instance, Hattori describes /r/ as a plosive stop when it precedes [i, ɯ] and a lateral before [a, e, o] (p. 98). Note that Kawakami (1977) also indicates these vowels as being amenable to laterality for Japanese in general. The trill [r], he says, can be heard in the *beranmee* (rough/highly casual) speech in the Edo dialect (p. 99), spoken by residents of Tokyo whose roots trace back to when modern-day Tokyo was pre-modern Yedo, or Edo. Of Kansai Japanese, he says that, in the phrase *kiko.en+rajio* (can:hear.NEG+radio; 'radio I can't hear'), /r/ is realized as a weak plosive (p. 56). Note that the other authors discussed here (e.g. Akamatsu 1997) indicate (of Standard Japanese) that post-/n/ is a likely environment for lateral realizations of /r/ to occur. Other than this disagreement over post-nasal /r/, Hattori is in agreement with the other authors in that he says intervocalic /r/ is a tap [ɾ] made with the underside of the tongue tip and the gums immediately behind the upper front teeth. Meanwhile word-initial /r/, or /r/ "that is not preceded by a vowel," is a weak plosive (p. 166).

2.6.5 Okada (1999)

Okada's view on /r/ appears in the International Phonetic Association's (1999) *Handbook*'s Japanese illustration. Okada suggests the symbol [ɽ], or retroflex tap/flap, be used to represent /r/ in Japanese but adds the proviso that, as opposed to retroflex, the Japanese sound is actually post-alveolar and occurs as such between vowels.

Following the velar (homorganic) nasal /ɴ/, Okada says that /r/ is realized as an affricate with short friction, for which he offers the symbol [d̠ɹ̆]. Without giving specific details, Okada also notes that the rhotic approximant [ɹ] may occur "in some environments" while a post-alveolar lateral approximant [l̠] is "not unusual in all positions" (p. 118).

2.6.6 Akamatsu (1997, 2000)

Akamatsu's view of /r/ is discussed in two volumes on the language: *Japanese Phonetics* (1997), and *Japanese Phonology* (2000). One aspect that makes Akamatsu's analysis of /r/ unique is that he describes two phonemes: a non-palatalized liquid /ɾ/ and a palatalized liquid /ɾ'/ (IPA: [$ɾ^j$]) (Akamatsu 2000:85). Akamatsu agrees with the other authors with respect to there being a difference between post-pausal and intervocalic /r/. He describes the post-pausal /ɾ/ "variant of [ɾ]" as starting out with the tongue tip loosely touching the ridge of the teeth then moving rapidly away (p. 106). Akamatsu suggests four possibilities with regard to laterals: palatalized and non-palatalized alveolar (lateral) taps ([ɺ'] and [ɺ]; IPA: [$ɺ^j$, ɺ]) as well as a post-[n] set (palatalized and not) of approximants: IPA [l^j, l].

2.7 *Articulatory studies of Japanese /r/*

This author's research uncovered no studies which looked at the articulation of /r/ in Kansai Japanese; however, three studies which include the articulatory properties of /r/ in a Tokyo, or 'standard' Japanese are available. These include Joo (2005), Kokken (1990), and Sudo et al. (1982). Joo (2005) and Kokken (1990) look at /r/ within the larger context of the entire phonemic inventory of Standard Japanese. Only Sudo et al. (1982) looks specifically at the articulation of /r/ as opposed to any other consonant sounds (including the palatalized 'variant of ɾ' which Akamatsu argues occurs phonemically before the vowels /a, ɯ, o/). All three studies mentioned here are based, moreover, on data elicited via reading lists using citation form as opposed to extemporaneous speech.

While the three studies use electro-palatography (EPG), Kokken (1990) additionally includes earlier cine-radiographic (X-ray film) tracings as well as measurements of air pressure within the vocal tract. Kokken's X-ray tracings and EPG results show a greater degree of tongue contact at the alveolar ridge for the post-pausal versus intervocalic contexts. Kokken's results also show greater palatal contact, and often full closure, for /r/ before [i] in both post-pausal and intervocalic contexts. They observe similar patterns for the phonemically palatalized 'variants of ɾ' before /a, ɯ, o/. With respect to these, often

Romanized as *rya, ryu, ryo,* Kokken (1990:493-494) argues that the glide component of the sounds do not follow the tongue-tip gesture but are rather coarticulated with it - a view shared by Akamatsu (1997). Kokken's EPG results show the least amount of palatal contact for the low, or retracted (Esling 2005) vowel /a/. Before /ɯ, e, o/, Kokken's EPG results show fewer contact points during /r/ than for /i/ or the coarticulated series, and they make no explicit mention of any laterals being produced in their dataset. Joo (2005) corroborates these observations, with the additional note that contact by the tongue along the palate proceeds from back to front.

Sudo et al. (1982) is a rare study in that it focuses entirely on how /r/ is articulated in Japanese. It presents an electro-palatographic study of intervocalic /r/ produced by two speakers of Tokyo Japanese. Stimuli consisted of V_1/r/V_2V_1/r/V_2 sequences which included all (25) combinations of the five vowels /i, e, ɯ, o, a/, produced in the frame sentence *'Sore wa ___ desu'* ('That is ___') .

Sudo and her colleagues found that there were two general patterns of linguo-palatal contact whereby their two subjects produced /r/. The most frequent pattern, in the environments of /i, e, ɯ/, showed contact between the tongue and palate proceeding from back to front along the alveolar ridge. In contrast, in the environments of /a/ and /o/, contact began at the anterior region of the palate. Their results further showed that /a, o/

environments least often featured complete closures between the tongue and palate while the opposite was true of pre- and post-/i/ environments. In other words, for Sudo et al.'s speakers, the vowels /a/ and /o/ involved more open articulations that proceeded from front-to-back while /i/, and to lesser extents /e, ɯ/, involved more complete closures that proceeded from back-to-front. They also report /a, o/ as showing the least amount of palatal contacts, with occasional lateral openings (Sudo et al. 1982:238).

Chapter Three

STUDY METHODOLOGY

3.1 The data

The data used in this study represents a small subset of a large corpus of spontaneous speech maintained by Dr. Nick Campbell and colleagues at Advanced Telecommunications Research Institute International (ATR), located in Kyoto, Japan. The larger corpus, the JST/ATR ESP-C corpus (see Campbell 2004, 2007 for full description), includes a longitudinal sampling of telephone-like conversations in Japanese between participants of varying ages. Participants, including both native and non-native speakers of Japanese, were paid to record ten sessions with each of the other participants. Other than filling the 30-minute block of time dedicated for each session, no prompting or 'speaking tasks' were assigned. The speech data analysed in the present thesis is comprised of two of those conversations: the tenth conversation between Japanese Female 'A' (JFA) and Japanese Female 'B' (JFB), and finally the tenth conversation between Japanese Male 'A' (JMA) and Japanese Male 'B' (JMB). Recordings from each pair's tenth session were chosen for the present study because it was felt that these would be most likely to reflect natural speech: the participants would not only be familiar and comfortable with the recording procedures but also with each other.

3.2 The speakers

Unfortunately, detailed language background information on each of the four speakers (JFA, JFB, JMA, JMB) in the two conversations was not available to this study, which means that biographical information about the four speakers had to be inferred from the recorded discourses themselves. Also, while the four talkers are native speakers of Japanese recorded in Kyoto (Kansai), they appear to differ as to the 'thickness' of their Kansai accents. By this author's impression, the two males (JMA, JMB) have the strongest Kansai-colouring to their speech while JFB's accent seems to most approximate the standard Tokyo dialect. JFA, for her part, speaks with an accent less strong than the two males but less standard than her conversation partner, JFB. The paragraphs that follow review the participants' biographical and demographic information inferred by this author during the data analysis process. Evaluations of the speakers' accents and demographic profiles by three native Japanese-speaking judges are also presented, in Section 3.3.

In terms of language background and other biographical information apparent in the recordings themselves, it is presumed that all speakers have had some exposure to English, at least through public education, in which it is mandatory. Speaker JMB (30's to 40's, author's judgement based on voice quality and discourse content) works as a full-time employee at a large hotel, which makes some degree of familiarity with foreign languages and highly reverential forms of standard Japanese quite likely. Speaker JMA (early 20's, author's judgement based on speech style and discourse content) works as a temporary part-time employee at a fast-food restaurant, among other short-term jobs (a so-called *furiitaa*, or 'free (part-time) hand'). It is unclear from the discourse whether JMA or JMB speak any language other than Japanese. Speaker JFA (50's, author's judgement based on voice quality and discourse content) is likely a homemaker, and has some exposure to English through study at English conversation school(s) which she makes passing reference to in the recording. No mention is made in the recording as to the duration or ongoing status of this as of the time the recording was made. Speaker JFB (20's, author's judgement based on speech style and discourse content) is a student of Chinese translation, which she studies at a private training college. JFB also has Australian relatives by marriage through a sibling, which suggests some degree of ongoing exposure or contact with Australian English. As indicated above, these are the author's impressions and interpretations taken from the content of both conversations, and not the judgments of native speakers of Japanese, which are presented in the following Section.

3.3 Native speaker accent evaluation

Since the principal investigator is not a native speaker of Japanese and therefore likely does not have as robust an intuition for ascertaining dialect and thickness of accent as would a native speaker, three native Japanese speakers (Japanese language teachers at the University of Victoria) were recruited to provide their impressions of the two conversations used here. The three judges (all female, non-Kansai speakers) were provided with a compact disc containing intensity-normalized digital recordings of both conversations as well as their respective transcriptions. An evaluation form (see Appendix 1) containing five questions (three open-ended; two Likert-type scale) constituted the judges' evaluation task. Although the form instructions indicated that the relative (i.e. older versus younger) ages of the speakers in each conversation, the first question on the evaluation form asked judges to estimate the age of each speaker. The second question on the evaluation form asked "What dialect of Japanese would you say each speaker is using?" Questions three and four asked judges to rate each speaker's strength of accent (Q. 3) and degree of formality (Q. 4) based on 5-point Likert-type scales. A rating of 1, for instance, meant a speaker used a very strong, non-Tokyo sounding accent for question 3 or very formal speech for question 4. The last question, Q. 5, was an open-ended re-casting of Q. 4 in that it also asked about relative formality. This time, judges were asked to choose which, if any, of the two speakers in either conversation spoke more formally than the other. Also, a small space was provided for judges to write why they thought a given speaker might be using more formal speech.

The rationale for asking judges to evaluate degree of formality and accent was to access sociolinguistic or pragmatic factors which might have a bearing on the way any given speaker spoke. That is, it was thought possible that the difference in how much 'Kansai-colouring' the speakers displayed may have been mitigated by the social context of their discourse. It is well known, for instance, that Japanese encodes certain social relationships between talkers in linguistically overt ways. Verbal affixes such as *-masu*, and honorific particles such as *o-* and *go-* are deployed to mark appropriate degrees of speaker humility and listener reverence. Similarly, lexical forms may differ depending on in-group/out-group relations, or what one might call social deixis. An example of this is the lexical distinction between one's own (i.e. in-group) family members and others' (out-group) family members: one's own mother as *haha* ('mother'), but someone else's mother as *okaasan* ('mother'). Given that Japanese discourse marks social deixis morphologically and lexically, it seems reasonable to expect that these relationships may be realized within the fine phonetic detail (Local 2003, 2007) as well. The role of judges' evaluation of the data, then, is to ascertain – albeit in general terms – what sort of social or pragmatic dynamics are being played out and thereby have a reference against which phonetic observations can be made.

The following Sections summarize the judges' assessments of each speaker analyzed in this study. As mentioned earlier, the evaluation task (Appendix 1) was general in focus and (perhaps unfortunately) not designed to ascertain what precise linguistic or phonetic cues may have informed judges' evaluations. Rather, it was assumed that the judges' evaluations would be informed individually based on their own experience as language instructors and as native speakers.

3.3.1 Judges' assessments: JFA

The three judges who evaluated the speakers examined here placed JFA's age in her 40s. One judge indicated '40s,' while the other two judges indicated 45 and 48 years, respectively. With respect to what dialect JFA used (Q. 2), responses were 'Osaka' (the largest urban prefecture in the Kansai region), 'Kansai,' and 'common Japanese.' Common Japanese, or *kyoutsuugo*, is a colloquial form of Standard Japanese which is not associated with any particular geographical area (Gottlieb 2005). In terms of strength of accent (Q. 3), which was rated on a 5-point Likert-type scale where '1' indicated Standard-like speech while '5' represented a very strong accent. The three judges gave JFA ratings of 1, 2, and 3 respectively, for an average of 2. A similar scale was used for Q. 4 which asked about the formality (1 = very formal; 5 = very informal) of each speaker's speech. The judges' ratings for JFA were 3, 2, and 4 (average 3). The last question that the judges responded to was which speaker in either pair they felt spoke more formally than the other and why. Two of the three judges responded that they felt difference in formality was likely due to the speakers' age difference, with JFB being the younger of the two. Oddly, however, these two judges disagreed as to which of JFA or JFB used more formal speech: one judge felt that JFA used less formal speech because she was older while another judge felt that JFA spoke more formally because she was older. The third judge felt neither JFA nor JFB spoke more formally than her partner, and that their speech was reflective of people who are acquainted with one another but not overly close.

3.3.2 Judges' assessments: JFB

Age-wise (Q. 1), two of the three judges felt that JFB was in her 20s, with one of whom specifically suggesting 28 years. The remaining judge placed JFB's age at 32 years. There was disagreement among the judges as to JFB's dialect (Q. 2): one judge indicated that, like JFA, JFB spoke *kyoutsuugo* - or common Japanese. Another judge responded that JFB's dialect was 'Kansai,' while the remaining judge indicated her dialect was Nagoya (neighbouring but not typically included among Kansai dialects). In terms of strength of accent (Q. 3) where a response of '1' indicated weak accent or most Standard-like speech and '5' the opposite, judges' ratings for JFB were 1, 2, and 3 (avg. 2). As for degree of formality (Q. 4: '1' = very formal, '5' = very informal), all three judges rated JFB's speech as '2'. As mentioned in the discussion of the judges' assessments of JFA for the last question, Q. 5, one judge felt that neither JFA nor JFB spoke more formally than her partner while the remaining two judges differed as to who they felt was more formal - one indicating JFA and the other JFB. Both of these judges cited age as the reason for the difference in formality.

3.3.3 Judges' assessments: JMA

JMA's age (Q. 1) was assessed as '20s' by one judge, and 22 and 26 years by the other two judges respectively. Two of the judges indicated for Q. 2 that he spoke the 'Kansai' dialect, while one judge more specifically indicated 'Osaka.' With respect to strength of accent (Q. 3), the judges' assessments were 3, 4, and 5 (avg. 4) where '5' indicated a 'very strong,' non-Standard accent. All three judges rated JMA's degree of formality (Q. 4) as the mid-range '3' – neither very formal nor very informal. As to which of JMA or JMB spoke more formally than the other (Q. 5), two of the three judges indicated JMA was slightly more formal. One of these two judges suggested that JMA used polite verbal inflections more often; the other judge indicated that JMA's greater degree or formality may be attributable to "his age or position at work."

3.3.4 Judges' assessments: JMB

The three judges placed JMB in an age range (Q. 1) similar to JFB: one indicated '20s,' another '21,' and the other '30.' As to JMB's dialect (Q. 2), the judges' responses were the same as for JMA: two judges indicated 'Kansai' while one specified 'Osaka.' Responses in regard to strength of accent (Q. 3) were 4, 4, and 5 (avg. 4.3) where, again, '5' indicates a very strong non-Standard accent. Judges' assessments of JMB's degree of formality (Q. 4) varied; these ratings were 3, 4, and 5 (avg. 4) where '5' indicates 'very informal.' One of the three judges felt that JMB spoke slightly more formally than JMA; however, the judge indicated uncertainty as to precisely why JMB spoke more formally.

3.4 Data preparation

The natural speech data made available to this study consists of two approximately thirty-minute conversations, a subset of the JST/ATR ESP-C corpus (Campbell 2004, 2007), between two pairs of participants who spoke with each other from separate sound-insulated rooms using headset microphones. The signal from each speaker's microphone was recorded as a separate digital audio file, thereby allowing speech analysis to be done on each talker's speech in isolation from that of their partner. The corpus used here, then, consists of four digital audio files: two from each conversation. Two additional files – each a synchronized combination of the two audio tracks in either conversation – were also made available to this study to allow for analysis of discourse features in context.

Using the digital audio editing software Audacity (Sourceforge.net 2006) as well as each conversation's orthographic transcription (included with the corpus), tokens of /r/ were extracted from individual speaker's recordings and saved with a unique file name in (MS) WAV format. Care was taken to include not only the word where the target /r/ was found, but the surrounding utterance as well. Where there occurred additional /r/s in the utterance, duplicate copies of the audio file were made and distinguished by a number to indicate if the /r/ in question was the 1st, 2nd, et cetera, in the utterance. In addition to this information, file names were assigned so as to reflect speaker identity (JFA, JFB, JMA, JMB), time index of the token within the recording, and a Roman-orthographic representation of the token's lexical context. In all a total of 1,535 audio files were created, each representing an individual token of /r/.

Lastly, it became apparent during initial examination of the different speaker-specific audio files that they varied as to how strong their acoustic signals were. For instance, with the two female speakers, JFA's recording was much louder than JFB's. In order to avoid this difference biasing the auditory analysis, all four speakers' token files were normalized for amplitude to 90% of maximum dB using an acoustics batch-processing software program, The WavNormalizer (LinearTeam 2000).

3.4.1 Segmentation

Using the speech analysis software Praat (Boersma & Weenink 2006), a Textgrid file was created for each token included in the acoustic analysis. The Textgrid function in Praat superimposes multiple annotation tiers onto sound spectrograms generated from digital sound files, like those used in the present study. These annotation tiers can then be used to define boundaries and add textual information. One such tier were used in this study to mark the boundaries of the r-domain, discussed below.

3.4.2 Defining an r-domain

In a given span of any sort of speech, identifying a single precise location where one speech sound begins and another ends is not a clear-cut proposition. Particularly in extemporaneous or natural conversation, which is replete with a great diversity of dysfluencies, reductions, and coarticulations, the task is a daunting one. Assuming an object of analysis which may take a variety of shapes, the present study opted to first define a domain (hereafter 'r-domain') within the speech signal large enough such that any given realization of /r/ could be captured. With this in mind, amplitude (a.k.a. intensity) was chosen as the primary indicator of the domain. Since vowels have higher amplitudes than consonants, and following from segmentation procedures in Warner & Arai (2001) and Warner et al. (2004), the r-domain was defined as the region between vocalic amplitude peaks. Figure 6 below illustrates the segmentation of the r-domain in one token of /r/ in the word *chiri* ('geography') spoken by speaker JFA.

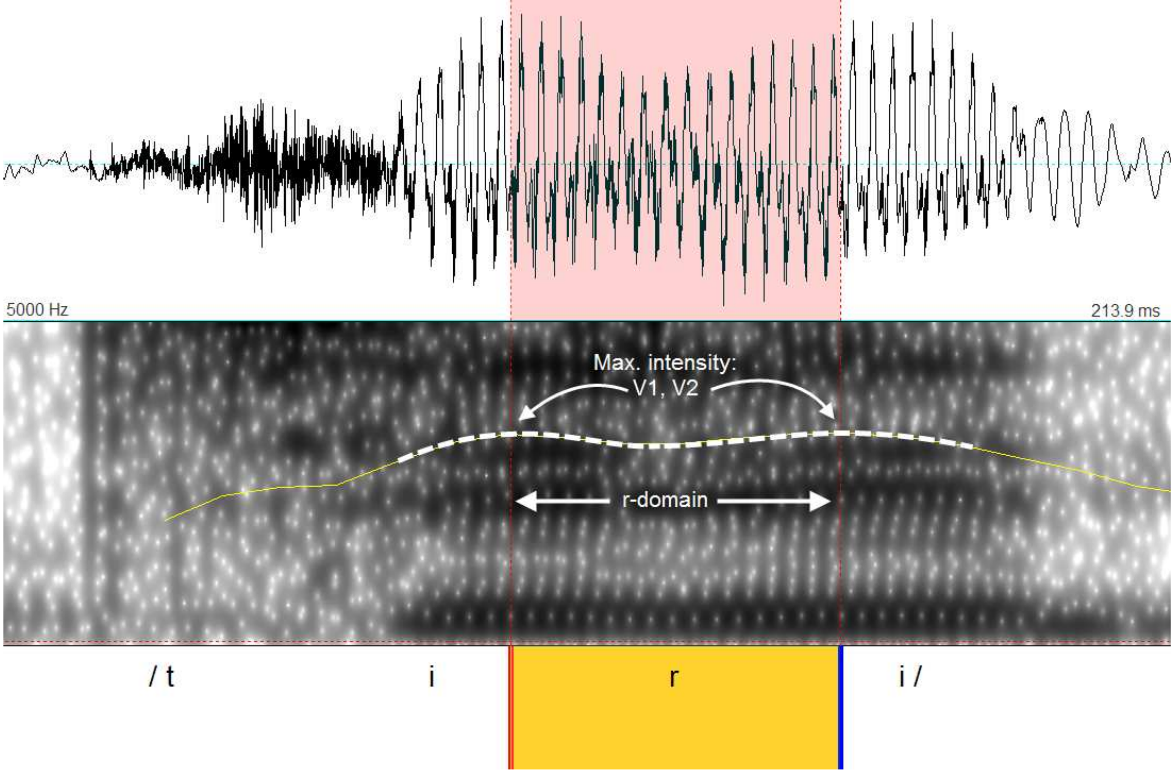


Figure 6. Segmentation of r-domain.

In the Figure, the thin grey line that indicates intensity in Praat has been highlighted by a bold dashed line. The r-domain is defined as the maxima of this dashed line for the vowels on either side of /r/.

3.5 Data transcription

With the 1,535 tokens of /r/ having been normalized for the intensity of their acoustic signal (amplitude), each token file was listened to repeatedly using headphones then narrowly transcribed into the International Phonetic Alphabet (IPA 2005). Care was taken to transcribe as many elements of the signal as possible, including voice quality (i.e. creak, breathiness, and nasality) as well as effects occurring between segments such as the rhoticization of vowels (V) and rhotic, palatal, or lateral off-glides from consonants (C). In terms of each transcription of /r/ being also a unit to be later coded into statistical analysis software, /r/ was defined as including these inter-segmental effects. That is, for instance, if a vowel preceding /r/ had a rhotic quality to it, or there was a rhotic or palatal off-glide into the following vowel, these were interpreted as being parts of the /r/ token and coded as such. In contrast to these inter-segmental effects, voice quality features (creak, breathiness, nasality) were interpreted as being supra-segmental and therefore not included as elements of /r/, but instead as diacritical elements of the surrounding vowels. In other words, the elements transcribed as belonging to a given token of /r/ included transitional effects from V to C to V, but voice quality was transcribed as a feature of V only. A complete listing of the symbols and diacritics used in the transcriptions is given in Table 3.

Table 3. IPA symbols and diacritics used in the narrow transcriptions.

IPA symbol	Explanation
(vowel diacritics)	
V̤	Breathy voice
V̰	Creaky voice
V̥	Devoicing
Ṽ	Nasalization
V̆	Shortness
Vː	Lengthening
(V→C transitions)	
V˞	Rhoticity, [ɹ]-like colouring
(C→V transitions)	
Cˡ	Lateral release
Cʲ	Palatal off-glide
Cʴ	Rhotic off-glide
(main consonant)	
d	Voiced alveolar stop
ɾ	Voiced alveolar flap
l	Voiced alveolar lateral approximant
ɺ	Voiced alveolar lateral flap
ɹ	Voiced alveolar rhotic approximant
(consonant diacritics)	
C̞	Lowering
C̝	Raising
C̥	Devoicing; frication noise
C̆	Shortness
C̃	Nasalization

It bears mention that the process of phonetic transcription was, at its most essential level, a subjective but instructed one wherein the primary instrument of analysis was the transcriber's (the author's) perception of each token of /r/ in the dataset. Also, while spectrograms of the speech signal were used to isolate the target /r/s, the transcriptions remain as faithful as possible to the auditory impression of each token as opposed to what its spectrogram may have 'shown.' That is, in short, where there seemed to be a conflict between auditory perception and the visual representation of the signal, the auditory judgement was given precedence. Given the inherent subjectivity involved in relying on the ears for the transcriptions, and not the eyes, it was necessary at the outset to establish operational criteria for choosing a given symbol over another. Especially for terms such as *rhotic*, which lacks a definitive phonetic definition (Lindau 1985), it was necessary to decide on what auditory criteria were to be met before the transcription process began. The sections that follow provide the operational definitions used in transcribing the data.

3.5.1 Transcribing rhoticity and laterality based on auditory impressions

Whether a given token of /r/ was judged as one of either rhotic, lateral, or neither rhotic nor lateral, was determined primarily by subjective analogy to the author's first language (L1) sound system. The author's L1 being a rhotic variety of Canadian English

(CE) with a contrast between the rhotic and lateral approximants [l] and [ɹ], an auditory colouring of either shade was relatively easy to detect. To clarify what is meant by the term *rhotic* with respect to the transcriptions done here, a token was considered rhotic if it had, at any time during its articulation, an auditory quality similar to that associated with the lowering or raising of the third formant frequency (F3) in North American English (e.g. Lindau 1985). Taking the English words *star* ([stɑ˞] in CE) and *ring* ([ɹĩŋ]) as illustrations, the last sound in *star* is a rhoticized vowel while the vowel in a word like *staw* ([stɑ]) is not rhotic in CE. Similarly, the sound leading into the vocalic portion of *ring* is rhotic, while there is no rhoticity in the portion orthographically represented by *-ing*. The cases of *star* and *ring* illustrate that rhoticity, in the way it is used in the present study, may be realized as a speech signal quality that can both follow and precede a vocalic interval. When following a vowel, there is the auditory impression of a declination in the signal (as in *star*), while the opposite is the case when the rhotic element precedes a vocalic interval (as in *ring*). Intervocalically, such as in the word *starring* ([ˈstɑɹĩŋ]), these conditions co-occur to give the hearer an auditory impression of a ‘dipping-then-rising’ between the vocalic sounds on either side.

The conventions used for transcribing rhoticity in the present study are as follows. For intervocalic tokens of /r/, where there was no audible interruption in resonance between the vocalic intervals but where there was a distinct 'dipping-then-rising' auditory quality similar to that in the CE word *starring*, the /r/ was transcribed using the symbol [ɹ]. The same symbol was used in post-pausal contexts where /r/ occurred initially, and there was a distinct auditory quality of rising intensity and third formant (F3) similar to that in the CE word *ring*. This symbol was not used, however, in cases where there could be heard any interruption in resonance (in intervocalic contexts), or where there occurred an audible release-like burst of sound into the following vowel. In these instances, where applicable, the symbol [˞] was used to indicate rhoticity from the preceding vowel into the consonant while the symbol [ʴ] was used to indicate rhoticity from the consonant into the following vowel. Finally, in intervocalic cases where there was no audible interruption in resonance between the vowels and only a faint 'dipping-then-rising' quality akin to but very much weaker than for [ɹ], the transcription [ɾ̞] (lowered flap) was adopted.

Analogous transcription conventions were adopted for transcribing laterality. A token was considered as having a lateral component if, at any point during its articulation, there was the audible raising or lowering signal similar to that associated with a rising or falling second formant (F2) in the English words *pil*, *low*, and *pillow*. When, in intervocalic contexts, a token's auditory signal had a lateral quality and also featured a brief interruption akin to that associated with non-lateral flaps, the symbol [ɺ] was used. For tokens which featured a strong, though brief, impression of laterality but no interruption in the auditory signal, the symbol [l̆] was used. Lastly, if a token featured no interruption in the auditory signal but a weak impression of laterality such that no central oral contact between the tongue and the alveolar ridge could be discerned, the symbol [l̞] was used.

3.6 Categorization

During the transcription process it soon became apparent that, due to the narrowness of the phonetic transcriptions, there were a wide variety of different – albeit closely related – realizations for /r/ in the dataset. For instance, tokens transcribed as raised flaps ([ɾ̝]) featured a greater interruption of surrounding vowels' resonance in a manner akin to, but less than, release bursts associated with voiced stops. Compared to tokens transcribed as [d], which were smaller in number, raised flaps seemed shorter in duration as well. Nonetheless, both of these types of /r/ were quite similar in that they could be characterized as being more similar to plosives than approximants. Further examples are tokens transcribed as rhotic approximants ([ɹ]) and lowered flaps ([ɾ̞]). Both of these involve little in the way of a central articulatory closure, and it so happened that the lowered flaps also had a rhotic, or [ɹ]-like quality. Similarly for tokens which had an [l]-like auditory quality, some of these involved a brief interruption in the resonance of the surrounding vowels - implying a momentary central contact in the oral vocal tract, for example, as in that involved in lateral flaps ([ɺ]). Other laterals featured no interruption in resonance, but sounded either sharply lateral (ex. [l̆]) or only slightly lateral and akin to vocoids in that there was no impression of a central articulatory contact at all (ex. [ɺ̞]).

The multitude of closely-related transcriptions made it necessary to group the various transcriptions into categories, which are summarized in **Table 4** below.

Table 4. Categorization of narrow phonetic transcriptions. Symbols preceded by '*' represent transcriptions which occurred more than 100 times in the dataset.

← *Central Oral Strictures* →

(No audible oral gesture)	**Neither Rhotic nor Lateral**	**Rhotic**	**Rhotic + Lateral**	**Lateral**
Strong/ Narrow	(ex. d, *ɾ̝, ɾ̝ʲ) release-like burst and/or frication	(ex. ˞ɾ̝, ˞ɾ̝ʴ, ɾ̝ʴ) rhoticity preceding or following a release-like burst and/or frication	(ex. ˞l, ˞l̆, ˞l̆ʲ) rhoticity preceding release-like burst and/or frication, with laterality into following vowel	(ex. l, *l̆, l̆ʲ) strong laterality with release-like burst/frication
Mid-Range	(ex. *ɾ, ɾʲ) brief interruption in audible resonance without frication or release-like burst	(ex. ˞ɾ, ˞ɾʴ, ɾʴ) rhoticity preceding or following brief interruption in resonance without frication or release-like burst	(ex. ˞ɺ) rhoticity preceding brief interruption in resonance without frication or release-like burst, but with laterality into following vowel	(ex. *ɺ, ɺʲ) brief interruption in audible resonance with laterality into following vowel
Weak/ Open	(ex. n/a) no interruption in resonance; no audible rhoticity or laterality	(ex. *ɹ, *ɾ̞) rhoticity with no interruption in audible resonance	(ex. ˞ɺ̞) no interruption in resonance, but rhoticity initially followed by laterality	(ex. l̞, *ɺ̞, ɺ̞ʲ) weak laterality with no interruption in resonance

← *Rhoticity vs. Laterality* →

Two general dimensions were decided upon for the categorization: 1) the auditory perception of the central articulatory stricture, and 2) the presence or absence of laterality and/or rhoticity. In Table 4, central oral stricture is given on the *y* axis while the laterality-versus-rhoticity dimension is on the *x* axis. Each cell lists the specific auditory criteria used for categorizing a variant as a member of that cell.

While the terms *narrow* and *open* are easily comprehensible insofar as they both relate to the physical space between the active and passive articulators (i.e. tongue and alveolar region), the accompanying terms *strong* and *weak* are much more problematic. That is, what exactly makes one articulation strong and another weak? Harris (2006:134) mentions that strength of articulation is an inadequately defined area that is nonetheless important in the production of (in particular) stops, fricatives, nasals, and trills. Citing long versus short consonants in Pattani Malay, Harris goes on to note that long consonants in this language use “more muscular strength, more compression, and tighter constriction than short consonants.” To the extent that flaps and taps can occur allophonically as reductions of both stops (as in North American English) and trills (as in various dialects of Spanish), the notion of strength of articulation is applicable to the study of Japanese /r/s as well. The question remains, however, as to how best to represent ‘strength’ as a phonetic variable.

The approach taken to this question in the present study is admittedly a surface one; *stricture* depicted in Table 4 is limited to the front of the tongue, and not the sides or root of the tongue. Also, stricture is divided into only three categories: those perceived as strong or narrow, those perceived as weak or open, and those perceived as neither narrow/strong nor open/weak (i.e. mid-range). In this way, a strong articulation in this study is restricted to the tightness of constriction at the front of the oral tract; muscular strength in open strictures is implied but unfortunately not measured *per se* aside from distinguishing them by how they were transcribed, as in the two open rhotics [ɾ̞, ɹ] where [ɹ] features a more robust auditory impression of rhoticity than does [ɾ̞].

3.7 Intra-rater reliability of transcriptions and categories

All of the transcription and categorization of the 1,535 tokens of /r/ was done by one individual, the author. Approximately one month following the completion of the transcriptions, an order to assess their reliability, 10% of the tokens were randomly selected, re-transcribed, and compared with their initial transcriptions. A re-transcription was judged as completely accurate only if there was total correspondence with its initial transcription. Recall that up to five elements were involved in transcribing any token of /r/: the main IPA symbol, combinatory diacritics above and below the main symbol, and diacritics representing transitions between the main symbol and adjacent vowels. The rate of complete accuracy of the re-transcriptions, where there was total correspondence of all five transcription elements, was 59.7%. Where there were differences between the initial transcriptions and re-transcriptions, these typically involved one diacritic, as in the difference between a non-lateral flap [ɾ] and a raised non-lateral flap [ɾ̝]. In a case such as this, the re-transcription was judged as inaccurate even though, in terms of the auditory-perceptual categories used here, the two transcriptions represented neighbouring categories: mid-range and strong/narrow non-rhotic non-laterals.

To the extent that a strict interpretation of transcription accuracy cannot distinguish between re-transcriptions that were quite close in agreement with their initial counterparts and those which were quite different, accuracy was also calculated based on a 5-point 'score.' That is, if each transcription reflected five elements, each of these could be interpreted as being consistent or not, and a score out of 5 could be calculated. The average score for the re-transcriptions was 4.3, where a score of 5.0 indicated complete correspondence with the initial transcriptions.

In terms of reliability across the categorical space defined by this study (see previous Section), reliability was assessed by calculating the distance separating the cell of an initial transcription and that of its re-transcription. To illustrate with the example of mid-range versus raised non-lateral flaps ([ɾ, ɾ̝]), both are instances of the non-rhotic non-lateral sounds; thus, along the laterality-versus-rhoticity dimension of the categorical space, there is no difference between the two transcriptions. However, along the stricture type dimension, a raised flap is an instance of the strong/narrow category whereas the plain flap is an instance of the mid-range category: the difference between which is 1 cell in the categorical space defined here. The distance between the two transcriptions would therefore be calculated as 1 cell along the stricture type dimension plus 0 cells along the laterality-versus-rhoticity dimension, for a total categorical space difference of 1. In cases where an initial transcription and its re-transcription differed along both the stricture type and laterality-versus-rhoticity dimensions, the distance between the two along both dimensions was combined. Overall, the average distance between the re-transcribed tokens and their initial counterparts was 0.54, or just over one half of one categorical space of distance.

Chapter Four

RESULTS

4.1 Transcription frequencies: Entire dataset

This section describes the various transcriptions for /r/ in the entire dataset (1,535 tokens). Table 5 below is a listing of those transcriptions which comprised 1.0% or more ($n \geq 15$) of the dataset.

Table 5. Transcription frequencies, entire dataset ($n \geq 15$, 1.0%)

Symbol	Description	Frequency	Percent
ɾ̝	raised alveolar non-lateral flap	218	14.2
ɹ	alveolar rhotic approximant	206	13.4
ɾ̞	lowered alveolar non-lateral flap	166	10.8
ɾ	alveolar non-lateral flap	161	10.5
ɺ̞	lowered alveolar lateral flap	142	9.3
ɺ	alveolar lateral flap	142	9.3
l̆	short alveolar lateral approximant	134	8.7
ɾ̥	tapped alveolar fricative	63	4.1
∅	deletion between same underlying vowels	36	2.3
l	alveolar lateral approximant	32	2.1
Vʲ	deletion for diphthong to [j]	31	2.0
ɾʴ	alveolar flap w/rhotic off-glide	30	2.0
˞ɺ	alveolar lateral flap w/rhotic onset	19	1.2
Vᵉ	deletion for diphthong to [e]	15	1.0
	Others ($n = < 15$)	140	9.1
		***T*= 1,535**	**100%**

The results in Table 5 suggest that /r/ was most commonly realized as a weak plosive, transcribed here as a raised alveolar flap ([ɾ̝]), which was observed 218 times in the entire dataset, accounting for 14.2% of the 1,535 /r/ tokens. This preference in the dataset for raised non-lateral flaps is even more apparent when these are combined with the 63 cases where additional noise leading into the following vowel was perceived as frication noise and not a release-like burst as for raised non-lateral flaps. For these, following Jesus & Shadle (2005) who observed voiceless tapped fricatives in European Portuguese, the symbol [ɾ̥] was adopted. Taken together, these two transcriptions ([ɾ̝, ɾ̥]) both categorized as strong/narrow non-rhotic non-laterals, represent 281 tokens, or 18.3% of the dataset.

The next most common transcription for /r/ was [ɹ], or the alveolar rhotic approximant ($n = 206$; 13.4%): the same speech sound as is found in North American English. Analogous to how raised non-lateral flaps and tapped alveolar fricatives were closely related to one another, rhotic approximants sounded similar to what were transcribed as lowered alveolar flaps ([ɾ̞]: $n = 166$; 10.8%). Specifically, lowered alveolar flaps had an auditorily rhotic, or [ɹ]-like quality to them but not to the extent observed for [ɹ]-proper. Taken together, these two transcriptions accounted for 24.2% of the entire dataset, or 372 tokens.

4.2 Descriptions of most frequent variants

This section further describes, from strong/narrow to weak/open stricture types, the variants which occurred most frequently in the dataset. While the intent of this study is not to rigorously elucidate the acoustic properties of /r/ realizations in Japanese, some rudimentary acoustical measurements as well as representative spectrograms will be referred to here to contextualize frequency observations discussed in later sections of this thesis. More detailed acoustical investigations of the compendium of /r/ sounds in this dataset remain a topic for future work.

4.2.1 Raised non-rhotic non-lateral flap [ɾ̝]

The auditory criteria for transcribing a realization of /r/ as [ɾ̝] hinged on the disruption of the vocalic sound signal or the presence of an extremely brief additional sound leading into the subsequent vowel. This brief additional sound was akin to but less prominent than the release burst associated with the voiced plosive [d]. Similarly, the overall disruption of the auditory signal associated with variants transcribed as [ɾ̝] was less prominent than that of [d]. Figure 7 below presents a representative spectrogram for one instance of a raised non-lateral flap spoken by JFB in the word *atarashiku* ('newly').

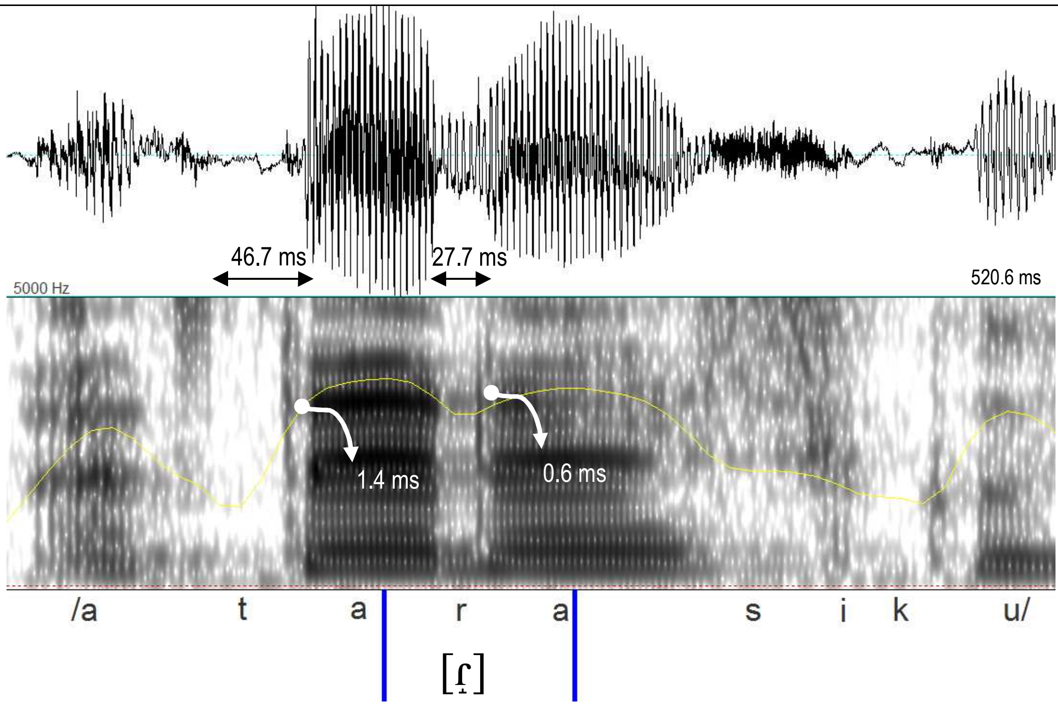


Figure 7. Spectrogram for raised non-lateral flap in *atarashiku* ('newly')

What is worthy of note in the spectrogram is the contrast between /t/ and /r/: both involve articulatory closures marked by a sharp dampening of sound energy, visible in the spectrogram as lighter shading. While the closure for /t/ is 46.7 milliseconds (ms), that of /r/ is shorter at 27.7 ms. Both sounds also involve release phases, including bursts, prior to the onset of their following vocalic intervals. The release phase for /t/ is longer than that of /r/: 1.4 ms compared to 0.6 ms.

4.2.2 Short alveolar lateral approximant [l̆]

Tokens were transcribed as short lateral approximants if they featured an auditory quality very much like [l] while at the same time being shorter than the sound associated with Canadian English /l/. Figure 8 depicts a spectrogram of two kinds of short lateral approximants spoken by JMA in the phrase *orekara-wa* ('from me-TOP').

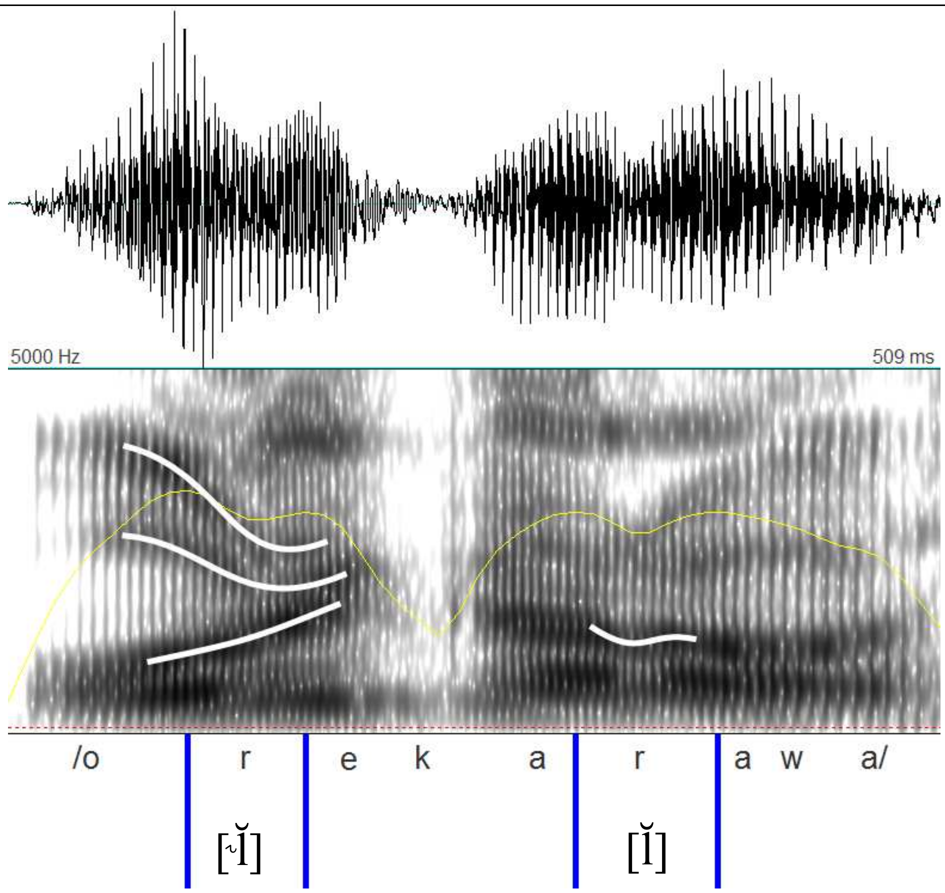


Figure 8. Spectrograms for two short lateral approximants in *orekara-wa* ('from me-TOP')

The first /r/ (transcribed as [˞l̆]) is a rhoticized short lateral where there was heard an [ɹ]-like quality leading into the consonant from its preceding vowel, /o/, and a strong but brief [l]-like quality leading from the consonant into the following vowel, /e/. Formants 2, 3, and 4 are highlighted by bold white lines in the Figure: F4 and F3 dip sharply while a very strong (dark in the spectrogram) F2 rises. The second /r/ (transcribed as [l̆], a short lateral approximant) also features a dark (high in energy) F2 which lowers-then-rises in concert with a slight decrease in amplitude.

4.2.3 Alveolar lateral flap [ɺ]

Tokens of /r/ were transcribed as lateral flaps if they involved an auditory quality akin to that of [l] that was also disrupted due to the momentary formation of an articulatory closure or near-closure along the centre line of the vocal tract. Figure 9 below is a spectrogram of two /r/s within a reduction/simplification of the phrase *dakara.ammar̠i* ('so not really,' realized as [ɣaɹa̰amːaɺi]) spoken by JMB. While the first /r/ is a rhotic approximant (discussed in Section 4.2.7), the second /r/ in the spectrogram represents a typical example of a lateral flap [ɺ].

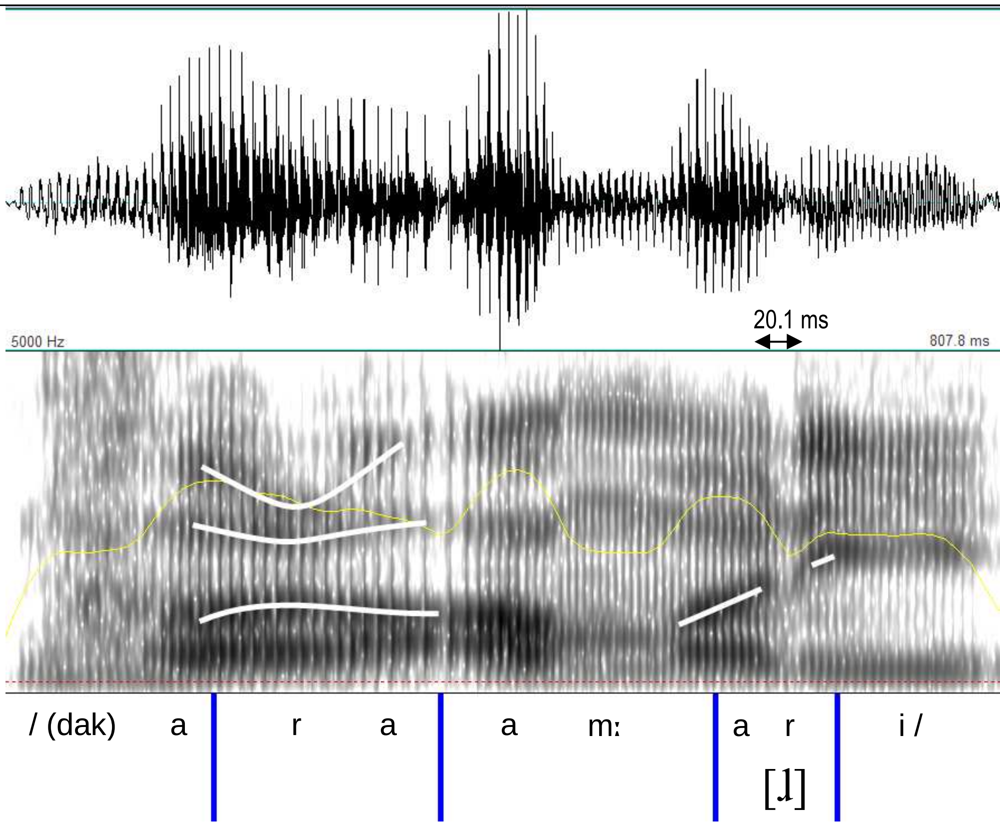


Figure 9. Spectrogram of a rhotic approximant [ɹ] and lateral flap [ɺ] in *dakara.ammari* ('so not really')

There is a brief (20.1 ms) period of decreased amplitude similar to that apparent in non-lateral flaps; however, different from raised non-lateral flaps, there is no burst-like release phase. This brief dip in amplitude, moreover, likely corresponds to what was perceived as an interruption in the vocalic speech signal - an auditory ‘bump in the road.’ The lateral, or [l]-like quality, likely corresponds to the behaviour of the second formant frequency (F2), highlighted by a bold white line on the spectrogram. As can be seen for [ɺ] in the Figure, F2 raises quite sharply, which is interpreted here as an acoustic cue indicative of laterality

(Ladefoged & Maddieson 1996). Whether the raising of F2 and its consequent impression of laterality is due to airflow around the sides of JMB's tongue in this case or, alternately, the impression of laterality is a function of the different F2 values of the adjacent vowels is not clear. To fully address this issue, instrumental techniques capable of visually detecting the presence or not of lateral airflow would have to be used.

4.2.4 Alveolar non-lateral flap [ɾ]

Tokens transcribed as non-lateral alveolar, or 'plain' flaps were those whose auditory signal involved a brief disruption but did not feature any quality reminiscent of [l]. Also, the signal disruption of non-lateral flaps did not have any additional burst-like release sound or audible frication noise preceding (i.e. leading into) the vocalic interval that followed the disruption. Figure 10 below shows a spectrogram of one instance of [ɾ] produced by JFB in the word *dakara* ('therefore').

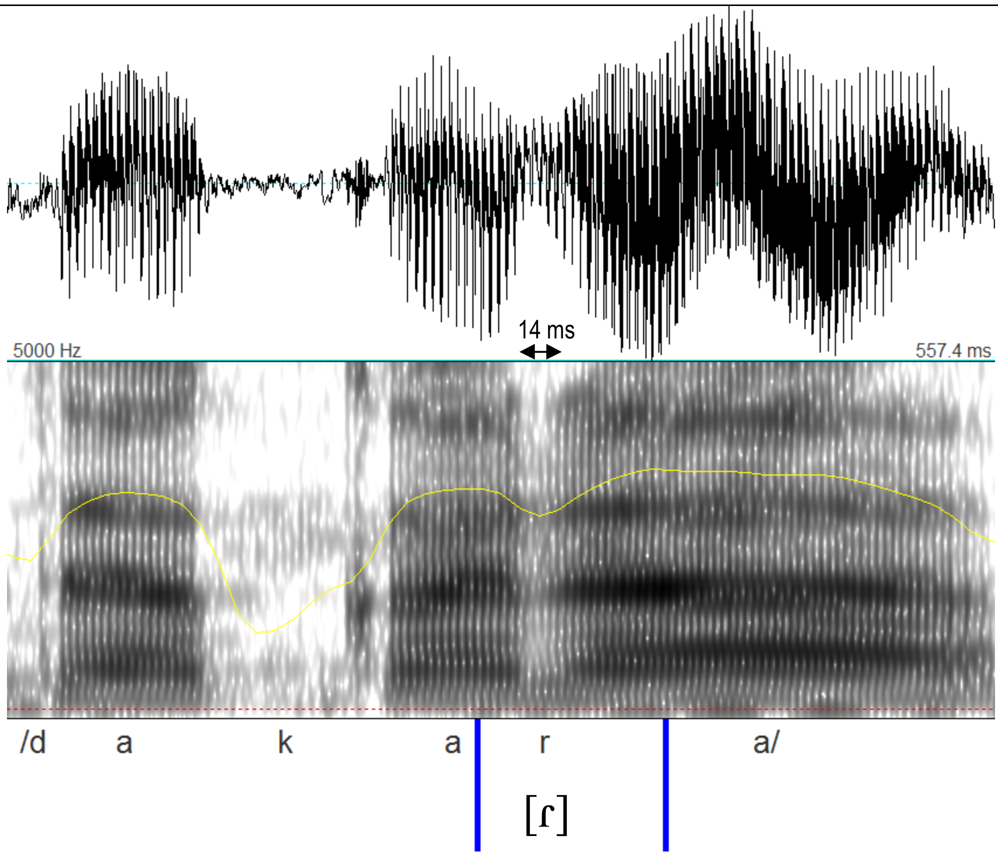


Figure 10. Spectrogram of alveolar non-lateral 'plain' flap in *dakara* ('therefore')

Amid what would otherwise be a prolonged /a/, a vertical band of lighter gray represents decreased energy (also shown by the thin intensity line dipping at the point of the vertical band). This interruption is quite brief - lasting only 14 milliseconds (ms). Unlike raised non-lateral flaps which entail a release-like burst of energy leading into the following vowel (see Figure 7), no such spike in energy is apparent here. Also, unlike lateral flaps whose second formants raise (see Figure 9), the second formant frequency here is stable.

4.2.5 Lowered alveolar lateral flap [l̞]

Tokens transcribed as lowered lateral flaps involved no interruption in the auditory signal but did feature a change in the quality of that signal akin to, but much weaker than, [l]. While it seems strange to talk about a lateral sound that has no central oral stricture, the presumption with regard to lowered laterals is that, during their production, lateral airflow exceeds central airflow to the point where an auditory quality resembling [l] is produced. The spectrogram in Figure 11 shows two instances of [l̞] spoken by JMA in a very rapid production of the word *gi<u>r</u>igi<u>r</u>i* ([gil̞ig̞il̞i], 'barely').

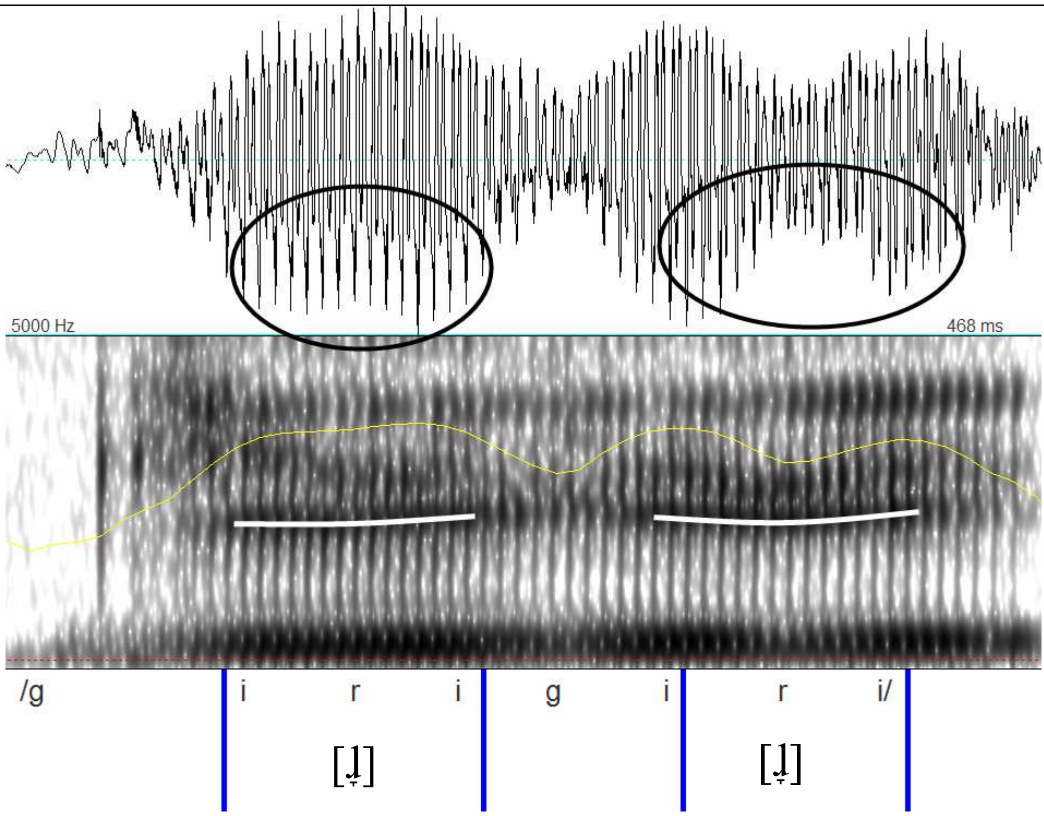


Figure 11. Spectrogram of lowered lateral flaps in *girigiri* ('barely')

Though more distinctive with the second token, both tokens involve a decrease in amplitude, indicated by black circles on the waveform and valleys in the thin gray line in the spectrogram. The bold white lines highlight the behaviour of F2, which raises slightly in both tokens.

4.2.6 Lowered non-lateral flap [ɾ̞]

Figure 12 shows a spectrogram of one instance of [ɾ̞] spoken by JFA in the word *dakar̲a* (‘therefore’).

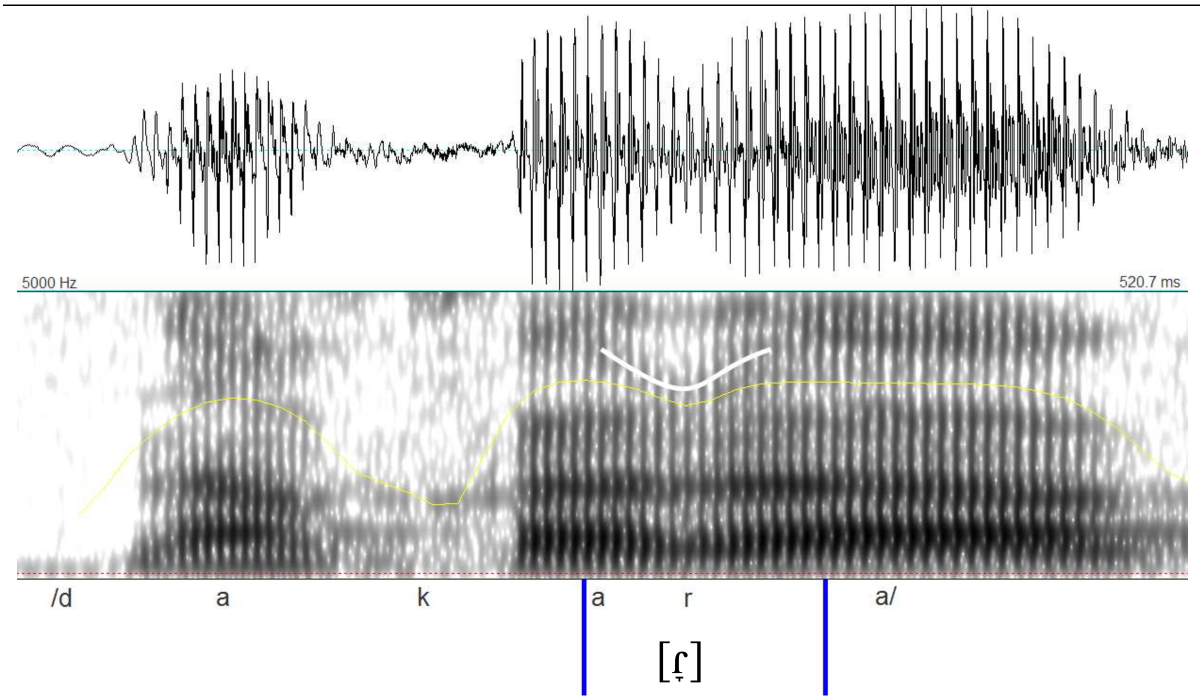


Figure 12. Spectrogram of lowered non-lateral flap in *dakara* ('therefore')

Tokens of /r/ were transcribed as lowered non-lateral flaps if their auditory signal featured neither any audible quality akin to [l], nor any interruption akin to raised or mid-range non-lateral flaps. These sounds did, however, present an auditory quality which was similar though less prominent than that of sounds transcribed as [ɹ], or rhotic approximants.

While it is clear from the waveform (upper pane in the Figure) that there is a decrease in amplitude during /r/, this decrease does not appear in the spectrogram (lower pane) as a distinctive vertical band of lighter shading, or decreased energy, which is interpreted here as a correlate of full or near complete obstruction along the centre of the vocal tract. In terms of formant frequency behaviour, the [ɾ̞] in Figure 12 shows a sharp dip in F4 (bold white line) much like the token of [ɹ] shown in Figure 13, presented in detail in the following Section. The third formant, however, is less obvious here than for [ɹ]: F3 seems to dissipate here as opposed to lowering. Other tokens of [ɾ̞] showed similar patterns, while still others had stable third formants but lowered fourth formants.

4.2.7 Alveolar rhotic approximant [ɹ]

Tokens were transcribed as [ɹ] if they involved no disruption of the auditory signal but rather a change in it akin to that involved in the middle of the Canadian English word *starring*. The sound represented by [ɹ] in North American English is typically associated with lowering of the third formant frequency (e.g. Delattre & Freeman 1968; Lindau 1985). While most (but not all) of the tokens of [ɹ] produced by the four speakers in this dataset involved F3 lowering, more consistent was lowering of the fourth formant (F4). The spectrogram below, repeated as Figure 13, again shows two instances of /r/ produced by JMB in a simplification ([ɣaɹa̰am:aɺi]) of the phrase *dakar̲a.ammari* (‘so not really’).

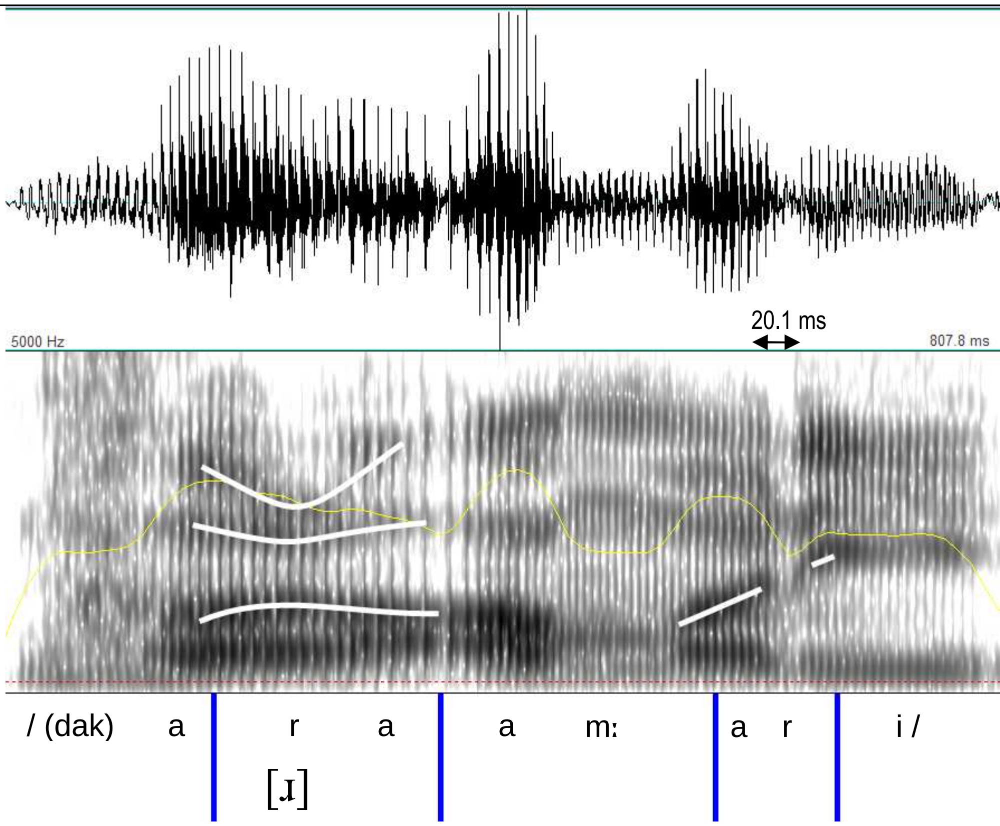


Figure 13. Spectrogram of a rhotic approximant [ɹ] and lateral flap [ɺ] in *dakar̲a.ammar̲i* ('so not really')

The first /r/ is a rhotic approximant [ɹ]: three bold white lines super-imposed on the spectrogram highlight (from top to bottom) formants 4, 3, and 2. Formants 3 and 4 both lower, approaching one another at their lowest point, roughly mid-way during /r/. At the same time, F2 raises toward the lowered F3. Unlike flaps which involve some disruption of the auditory signal through the formation of a brief articulatory closure or near-closure such as in the lateral flap in r(2), there is no such disruption involved in the production of [ɹ].

4.3 Amplitude as measure of auditory signal disruption

Amplitude, measured in decibels (dB), is the acoustic correlate of what the hearer of a sound might perceive as *loudness*, or *intensity*. The rationale for looking at the different transcriptions in terms of their amplitudes stems from the assumption that the intensities of different /r/ variants will be expressed in different ways. For instance, raised flaps were perceived as involving complete, if brief, articulatory closure and resultant 'interruption' in the auditory signal. This interruption, in turn, is arguably an instance of briefly lowered amplitude caused by the momentary obstruction of the vocal tract which occurs when the tongue comes close to or in contact with the alveolar ridge. The key assumption at work here is that obstruction of the oral vocal tract results in lowered amplitude when compared with vocalic intervals which involve little obstruction of the oral vocal tract; more obstruction, in other words, means less amplitude.

With the above assumption in mind, the next issue in characterizing /r/ variants by their amplitudes is how to go about doing it. This study's data in particular are a challenge in this regard in that the data represent different speakers whose productions have no external constraints placed on them in terms of what was to be said or how. Due to the data being taken from extemporaneous speech, that is, amplitudes of individual tokens of /r/ ranged widely. This is unsurprising if one considers how people of any language talk to one another in casual, every-day life: voices increase and decrease in volume to reflect not only linguistically salient chunks of information but also to show conformity with the pragmatics of the conversation going ahead. For this reason, the amplitude measurements described here are ratios.

Recall that the r-domain for intervocalic tokens of /r/ was defined as falling between the intensity peaks of the preceding and following vowels. Following Warner et al. (2004) and Tucker & Warner (2007), the average dB of the start and end of the domain was used as a reference point against which to compare the minimum amplitude that occurred within the domain. That is, averaging the start and end amplitudes provides an intensity 'ceiling,' from where we can say how close the floor is. The floor, in this case, is the minimum amplitude value within the r-domain.

Figure 14 shows, in order from greatest to least, the ratio (%) of each transcription's minimum amplitude to the averaged amplitude of the start and end of the r-domain. In other words, what proportion of the ceiling amplitude the minimum comprises. Higher proportions are interpreted as involving less obstruction of the oral vocal tract while lower proportions are interpreted as involving more obstruction of the oral vocal tract.

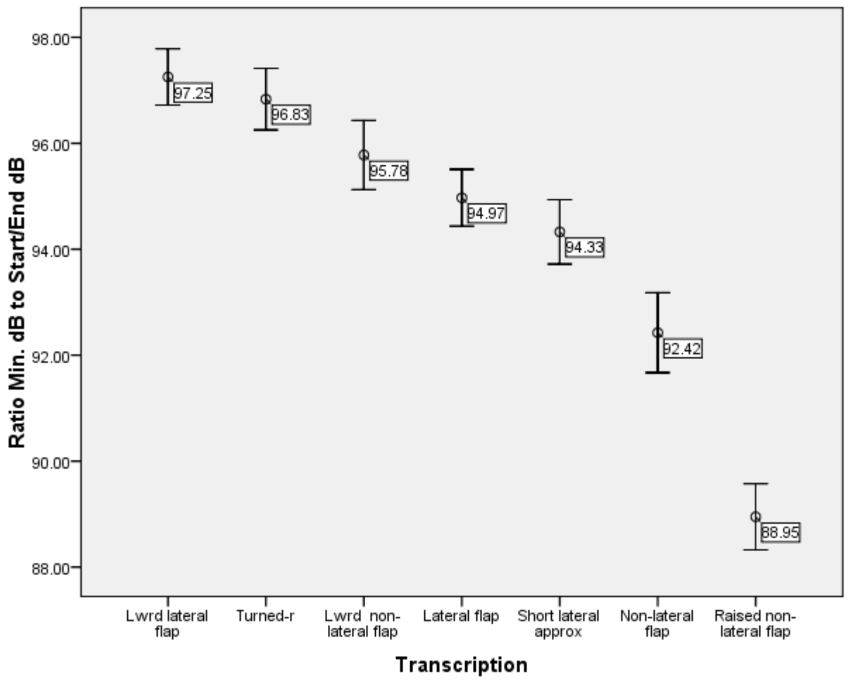


Figure 14. Ratio of minimum amplitude to averaged start and end amplitudes (whiskers indicate one standard error).

The results shown in Figure 14 suggest a correspondence between obstruction of the oral vocal tract and decreased amplitude. On average, the minimum dB of tokens transcribed as [ɺ̞], or lowered (open) lateral flaps, was 97.25% of the peaks of their surrounding vowels. For rhotic approximants [ɹ], this ratio was 96.83%. As for the five remaining transcriptions, the ratio was 95.78% for lowered non-lateral flaps [ɾ̞], 94.97% for lateral flaps [ɺ], and 94.33% for short lateral approximants [l̆]. Markedly lower ratios were apparent for non-lateral flaps [ɾ] at 92.42% and raised non-lateral flaps [ɾ̝] at 88.95%.

4.4 Category frequencies: Entire dataset

Due to the narrowness of the phonetic transcriptions involved in this study, it was necessary to formulate categories which grouped these transcriptions together so as to facilitate analysis. Categories (see Methodology) were established based on the perceived strength/narrowness of the central articulatory stricture as well as the perceived presence or absence of laterality and/or rhoticity. Category frequencies for the entire dataset are summarized in Table 6 as follows.

Table 6. Category frequencies, entire dataset, % (*n*/ 1,535 tokens)

← *Stricture* →		≠Rhotic ≠Lateral	Rhotic	Rhotic+ Lateral	Lateral
Strong/ Narrow	35.0% (537)	**20.7%** (317)	**1.1%** (17)	**1.0%** (15)	**12.2%** (188)
Mid- Range	31.9% (489)	**10.6%** (163)	**10.6%** (163)	**1.2%** (19)	**9.4%** (144)
Weak/ Open	34.7% (532)	--	**25%** (383)	**0.1%** (2)	**9.6%** (147)
No Oral Gesture	**6.1%** (94)	31.3% (480)	37.7% (563)	2.3% (36)	31.2% (479)
		← *Rhoticity vs. Laterality* →			

Stricture type in the Table is depicted along the *y* axis while laterality-versus-rhoticity appears on the *x* axis. Non-overlapping cells to the right of stricture type labels and above laterality-vs.-rhoticity labels represent totals for that label. Recall that weak/open non-lateral flaps ([ɾ̞]), which would presumably occupy the open, non-rhotic non-lateral cell in Table 6 were perceived as having an [ɹ]-like quality; for this reason, instances of [ɾ̞] were categorized as open rhotics.

Non-rhotic non-lateral variants with mid- (ex. [ɾ]) to strong/narrow strictures (ex. [d, ɾ̝]) accounted for roughly 30% of the data while rhotics with mid- (ex. [(V)˞ɾ, ɾʴ]) to weak/open strictures (ex. [ɹ, ɾ̞]) made up roughly 36%. Laterals comprised a further rough 30% of the dataset ranging from those with perceptually strong central strictures (ex. [l, l̆]) to weak/open strictures (ex. [ɺ̞]), as well as mid-range ones (ex. [ɺ]).

4.4.1 Category frequencies: Entire dataset, by speaker

Table 7, below, summarizes the category frequencies of each of this study's four speakers: JFA, JFB, JMA, and JMB. The table has four parts to it, each part representing the category matrix for each speaker as shown in the previous Section, with stricture types depicted on the *y* axis and laterality-versus-rhoticity along the *x* axis.

Table 7. Category frequencies for all speakers in all environments (S/N: strong/narrow, M: mid-range, W/O: weak/open; ≠RL: non-rhotic non-lateral, R: rhotic, R+L: rhotic plus lateral, L: lateral; NOG: no audible oral gesture)

JFA, % (*n*/315 tokens)

S/N	28.3 (89)	20.3 (64)	1.0 (3)	0.3 (1)	6.7 (21)
M	28.6 (90)	15.9 (50)	1.9 (6)	0.6 (2)	10.2 (32)
W/O	40.6 (128)	--	27.6 (87)	0.0 (0)	13.0 (41)
NOG	2.5 (8)	36.2 (114)	30.5 (96)	1.0 (3)	29.8 (94)
n =	315	≠RL	R	R+L	L

JMA, % (*n*/463 tokens)

S/N	22.2 (103)	1.5 (7)	0.2 (1)	1.5 (7)	19.0 (88)
M	24.0 (111)	5.2 (24)	5.0 (23)	2.4 (11)	11.4 (53)
W/O	42.3 (196)	--	32.4 (150)	0.0 (0)	9.9 (46)
NOG	11.4 (53)	6.7 (31)	37.6 (174)	3.9 (18)	40.4 (187)
n =	463	≠RL	R	R+L	L

JFB, % (*n*/381 tokens)

S/N	61.7 (235)	55.9 (213)	1.3 (5)	0.0 (0)	4.5 (17)
M	26.0 (99)	18.4 (70)	1.0 (4)	0.0 (0)	6.6 (25)
W/O	11.0 (42)	--	3.7 (14)	0.0 (0)	7.3 (28)
NOG	1.3 (5)	74.3 (283)	6.0 (23)	0.0 (0)	18.4 (70)
n =	381	≠RL	R	R+L	L

JMB, % (*n*/376 tokens)

S/N	29.3 (110)	8.8 (33)	2.1 (8)	1.9 (7)	16.5 (62)
M	19.1 (72)	5.1 (19)	3.5 (13)	1.6 (6)	9.0 (34)
W/O	44.1 (166)	--	35.1 (132)	0.5 (2)	8.5 (32)
NOG	7.4 (28)	13.8 (52)	40.7 (153)	4.0 (15)	34.0 (128)
n =	376	≠RL	R	R+L	L

It is worth restating that the four speakers were paired by gender in the two conversations examined here. That is, the two females (JFA, JFB) represent one conversation while the two males (JMA, JMB) represent the other. The two male speakers divided the bulk of their tokens largely between open/weak rhotics and strong/narrow laterals. One female speaker (JFB), in contrast, realized over 70% of her /r/'s as mid- to strong/narrow non-lateral non-rhotic articulations.

In the sections to follow, which display category results for various phonological environments in tables similar to Table 7, each speaker's percentage of a given category will be compared to their overall average listed above. Higher than average percentages will be indicated by shading (e.g. '%'), while lower than average percentages will be underlined (e.g. '%').

4.5 *Phonemic environment frequencies*

This section summarizes the phonemic environments where the 1,535 /r/ tokens occurred; Table 8 presents these environments in descending order of frequency in the entire dataset.

Table 8. Phonemic environment frequencies

Environment	Frequency	Percent	Environment	Frequency	Percent
a_a	281	18.3	e_e	20	1.3
o_e	179	11.7	ɯ_o	18	1.2
a_i	166	10.8	o_o	16	1.0
e_ɯ	132	8.6	ɯ_i	15	1.0
a_ɯ	117	7.6	o_ɯ	13	0.8
a_e	107	7.0	e_a	12	0.8
ɯ_a	85	5.5	e_i	8	0.5
ɯ_ɯ	47	3.1	#_e	5	0.3
i_o	46	3.0	#_a	4	0.3
i_a	40	2.6	#_o	4	0.3
o_a	39	2.5	#_i	3	0.2
o_i	39	2.5	e_o	2	0.1
ɯ_e	35	2.3	n_e	2	0.1
i_ɯ	30	2.0	#_ɯ	1	0.1
a_o	23	1.5	n_a	1	0.1
i_i	23	1.5	n_i	1	0.1
i_e	20	1.3	n_o	1	0.1
			Total:	**1,535**	**100%**

While /r/ occurred in all vocalic environments as well as post-pausally/word-initially, it occurred most frequently either after /a/ or before /ɯ/. Aside from purely lexical considerations, the high frequency of the most common environments is likely due to the fact that these are also employed by the language in certain grammatical functions. For instance, the most common environment /a_a/ (*n* = 281; 18.3%) is found in the conditional morphemes *–tara* and *–nara*, as well as *mizenkei* forms of verbs to which the negative affix *-nai* attaches. The environment /o_e/ (*n* = 179; 11.7%) is found in the demonstratives *kore* ('this') and *sore* ('that'), as well as the informal first-person singular pronoun *ore* ('I') used by males. /a_i/ (*n* = 166; 10.8%) appears in the morpheme *–tari* ('and so forth'), which is affixed to sequential verb phrases to convey the same sense as does falling intonation in an English sequence like, "I did the dishes, baked a cake, fed the cat, and took out the garbage."

The high frequencies of environments where /r/ occurs before /ɯ/ is due to the fact that the infinitival "dictionary" forms of a great many verbs and verbial inflections in Japanese end with the sequence *-ru*, as in *suru* ('do') and *aru* ('to exist'). The environment /e_ɯ/ (*n* = 132; 8.6%) often occurs as a colloquial simplification of the verbial inflection -*iru*, which is affixed to the connective/*renyoukei*, or "*-te*" form of verb stems to form the imperfective as in *shite-iru* ('doing'). In casual speech, this is commonly realized as *shite-ru.* The environment /e_ɯ/ is also found in passives, as in *sareru* ('be done').

Surprisingly infrequent in the dataset were post-pausal, and post-nasal tokens of /r/. The low frequency of post-pausal tokens is likely due to the fact that, in connected speech, word-initial /r/s are produced immediately following the last vowel of a preceding word. That is, even though a token of /r/ occurs at the beginning of a word, its close proximity to the word it follows makes it intervocalic. Post-nasal environments are limited to word-medial syllable boundaries, as in the word *sen.ro* ('track') where /n/ closes the syllable preceding the one opened by /r/.

The sections that follow show transcription and category frequencies for the intervocalic environments where /r/ occurred more than one hundred times, in order of frequency: /a_a, o_e, a_i, e_ɯ, a_ɯ, a_e/. From there, transcription and category frequencies for /r/ before and following the five vowels /i, ɯ, e, o, a/ are shown, followed by results for post-pausal and post-nasal environments. Finally, observations for /r/ before, during, and following vowels perceived as creaky are discussed.

4.6 Common intervocalic environments

4.6.1 /a_a/

Occurring 281 times in the dataset, the sequence /ara/ was by far the most common intervocalic environment where /r/ was observed. By itself, *ara* is an interjection ('Oh dear!'), and the sequence also occurs in very frequent functional situations such as in the postposition *k<u>ara</u>* ('from/because') and the conditional affix *-tara* (*nat-t<u>ara</u>* 'become-COND'). The sequence also frequently occurs in the *mizenkei* form of verbs, used to form negatives as in *n<u>ara</u>-nai* ('become-NEG').

The transcriptions for the 281 /r/s in the /a_a/ context are summarized in Table 9, and their categories are summarized in Table 10.

Table 9. Transcriptions for /r/ in /ara/ (NOG: no audible oral gesture). Values (*n*, %) apply additively to each symbol where more than is listed.

All speakers (*n*= 281; 100%)	**JFA (*n*= 54; 100%)**	**JFB (*n*= 57; 100%)**	**JMA (*n*= 83; 100%)**	**JMB (*n*= 87; 100%)**
ɹ (*n*= 47; 16.7%)	ɾ̞ (*n*= 14; 25.9%)	ɾ̝ (*n*= 25; 43.9%)	NOG (*n*= 18; 21.7%)	ɹ (*n*= 34; 39.1%)
l̆, NOG (*n/2*= 33; 11.7%)	ɺ (*n*= 9; 16.7%)	ɾ (*n*= 12; 21.1%)	l̆ (*n*= 15; 18.1%)	NOG (*n*= 11; 12.6%)
ɾ̞ (*n*= 32; 11.4%)	ɾ (*n*= 7; 13.0%)	ɺ, ɺ̞ (*n/2*= 5; 8.8%)	ɹ (*n*= 11; 13.3%)	l̆, ɺ̞ (*n/2*= 10; 11.5%)
ɺ, ɺ̞ (*n/2*= 28; 10.0%)	l̆, ɺ̞ (*n/2*= 6; 11.1%)		ɺ, ɾ̞ (*n/2*= 8; 9.6%)	ɾ̞ (*n*= 8; 9.2%)
ɾ̝ (*n*= 27; 9.6%)	l (*n*= 5; 9.3%)		ɺ̞ (*n*= 7; 8.4%)	ɺ (*n*= 6; 6.9%)
ɾ (*n*= 26; 9.3%)	NOG (*n*= 3; 5.6%)		ɾ (*n*= 6; 7.2%)	
others (*n*= 27; 9.6%)	others (*n*= 4; 7.4%)	others (*n*= 10; 17.5%)	others (*n*= 10; 12.0%)	others (*n*= 8; 9.2%)

Table 10. Category frequencies for /a_a/ (S/N: strong/narrow, M: mid-range, W/O: weak/open; ≠RL: non-rhotic non-lateral, R: rhotic, R+L: rhotic plus lateral, L: lateral; NOG: no audible oral gesture)

JFA, % (*n*/ 54 tokens)

	n = 54	≠RL	R	R+L	L
S/N	22.2 (12)	1.9 (1)	0.0 (0)	0.0 (0)	20.4 (11)
M	29.6 (16)	13.0 (7)	0.0 (0)	0.0 (0)	16.7 (9)
W/O	42.6 (23)	--	29.6 (16)	0.0 (0)	13.0 (7)
NOG	5.6 (3)	14.8 (8)	29.6 (16)	0.0 (0)	50.0 (27)

JMA, % (*n*/ 83 tokens)

	n = 83	≠RL	R	R+L	L
S/N	24.1 (20)	0.0 (0)	0.0 (0)	0.0 (0)	24.1 (20)
M	21.7 (18)	7.2 (6)	4.8 (4)	0.0 (0)	9.6 (8)
W/O	32.5 (27)	--	21.4 (20)	0.0 (0)	8.4 (7)
NOG	21.7 (18)	7.2 (6)	28.9 (24)	0.0 (0)	42.2 (35)

JFB, % (*n*/ 57 tokens)

	n = 57	≠RL	R	R+L	L
S/N	52.6 (30)	47.4 (27)	0.0 (0)	0.0 (0)	5.3 (3)
M	33.3 (19)	24.6 (14)	0.0 (0)	0.0 (0)	8.8 (5)
W/O	12.3 (7)	--	3.5 (2)	0.0 (0)	8.8 (5)
NOG	1.8 (1)	71.9 (41)	3.5 (2)	0.0 (0)	22.8 (13)

JMB, % (*n*/ 87 tokens)

	n = 87	≠RL	R	R+L	L
S/N	14.9 (13)	1.1 (1)	1.1 (1)	0.0 (0)	12.6 (11)
M	10.3 (9)	1.1 (1)	2.3 (2)	0.0 (0)	6.9 (6)
W/O	62.1 (54)	--	49.4 (43)	1.1 (1)	11.5 (10)
NOG	12.6 (11)	2.3 (2)	52.9 (46)	1.1 (1)	31.0 (27)

No single transcription was overwhelmingly preferred by all four speakers, although the two male speakers (and to a lesser extent speaker JFA) realized /r/ most frequently as either open rhotics ([ɾ̞, ɹ]), laterals with mid-range to narrow strictures ([l̆, ɺ]), or not at all (NOG: no audible oral gesture). For speaker JMA, deletions (NOG) were most frequent at 21.7% (*n*= 18) followed by the short lateral approximant ([l̆]: 18.1%, *n*= 15) and rhotic approximant ([ɹ]: 18.1%, *n*= 15). The rhotic approximant made up the bulk of

JMB's productions at 39.1% (*n*= 34) followed by deletions at 12.6% (*n*= 11) and 10 tokens (11.5%) each of the short lateral approximant and lowered (open) lateral flap [ɺ̞]. Speaker JFA, while having much fewer deletions than the two males (5.6%, *n*= 3), patterned similarly in that she showed an overall preference for productions categorized either as rhotic or lateral over non-lateral, non-rhotic sounds. She produced the lowered non-lateral flap ([ɾ̞], an open rhotic) at 25.9% (*n*= 14) followed by the lateral, then non-lateral flaps at 16.7 and 13% (*n*= 9, 7) respectively. For JFB's part, her realizations of /r/ in this context were consistent with her productions in most other phonological environments: she showed a strong preference for the raised (i.e. narrow/strong stricture) non-lateral flap [ɾ̝] (43.9%, *n*= 25) followed by the plain (i.e. mid-range stricture) non-lateral flap [ɾ] at 21.1% (*n*= 12). Despite this, JFB produced fewer raised non-lateral flaps here compared to her overall average; also, she produced more mid-range non-rhotic non-laterals as well as laterals of all stricture types.

4.6.2 /o_e/

Apart from non-functional lexical items, the sequence /ore/ appears in a variety of function words and verb conjugations – making its high frequency in the dataset unsurprising. Specific examples include the demonstrative pronouns *kore* ('this') and *sore* (‘that’), as well as one of two first-person pronouns normally used by men in casual speech: *ore* (‘I’). Of the two male speakers in this study, JMA preferred the pronoun *ore* to the other first-person pronoun associated with less-formal masculine speech, *boku.* The other male speaker, JMB, tended to use the latter of the two. As can be noticed in the distribution of tokens in Table 11, which summarizes the transcriptions for /r/ in the environment /o_e/, JMA’s substantially greater number of tokens than the other speakers is most likely due to his choice of pronouns. The distribution of the tokens’ perceptual categories is summarized in Table 12.

Table 11. Transcriptions for /r/ in /o_e/ (NOG: no audible oral gesture). Values (*n*, %) apply additively to each symbol where more than is listed.

All speakers (*n*= 179; 100%)	JFA (*n*= 25; 100%)	JFB (*n*= 41; 100%)	JMA (*n*= 75; 100%)	JMB (*n*= 38; 100%)
ɹ (*n*= 38; 21.2%) ɭ̟ (*n*= 22; 12.3%) NOG (*n*= 20; 11.2%) ɾ̟ (*n*= 15; 8.4%) ɾ̠ (*n*= 13; 7.3%) ɾ̥, ɾ, ɭ̆ (*n/3*= 12; 6.7%) ˞ɭ (*n*= 9; 5.0%)	ɹ, ɾ̟ (*n/2*= 6; 25.0%) ɾ̠, ɾ (*n/2*= 3; 12.0%) ɭ̟ (*n*= 2; 8.0%)	ɾ̠ (*n*= 10; 24.4%) ɾ̥ (*n*= 9; 22.0%) ɾ (*n*= 8; 19.5%) ɭ̟ (*n*= 5; 12.2%) NOG (*n*= 3; 7.3%)	ɹ (*n*= 25; 33.3%) NOG (*n*= 12; 16.0%) ˞ɭ (*n*= 7; 9.3%) ɾ̟ (*n*= 6; 8.0%) ˞ɭ̆, ɭ̆ (*n/2*= 5; 6.7%)	ɹ, ɭ̆ (*n/2*= 6; 15.8%) ɭ̟ (*n*= 5; 13.2%) NOG (*n*= 4; 10.5%) ˞ɭ̆, ɾ̥ (*n/2*= 3; 7.9%) ɭ, ˞ɾ, ɾ̟ (*n/3*= 2; 5.3%)
others (*n*= 26; 14.5%)	others (*n*= 5; 20.0%)	others (*n*= 6; 14.6%)	others (*n*= 15; 20.0%)	others (*n*= 5; 13.2%)

Table 12. Category frequencies for /o_e/ (S/N: strong/narrow, M: mid-range, W/O: weak/open; ≠RL: non-rhotic non-lateral, R: rhotic, R+L: rhotic plus lateral, L: lateral; NOG: no audible oral gesture)

JFA, % (*n*/ 25 tokens)

	n = 25	≠RL	R	R+L	L
S/N	20.0 (5)	16.0 (4)	0.0 (0)	0.0 (0)	4.0 (1)
M	20.0 (5)	12.0 (3)	4.0 (1)	4.0 (1)	0.0 (0)
W/O	56.0 (14)	--	48.0 (12)	0.0 (0)	8.0 (2)
NOG	4.0 (1)	28.0 (7)	52.0 (13)	4.0 (1)	12.0 (3)

JMA, % (*n*/ 75 tokens)

	n = 75	≠RL	R	R+L	L
S/N	14.7 (11)	0.0 (0)	1.3 (1)	6.7 (5)	6.7 (5)
M	13.3 (10)	0.0 (0)	0.0 (0)	9.3 (7)	4.0 (3)
W/O	56.0 (42)	--	42.7 (32)	0.0 (0)	13.3 (10)
NOG	16.0 (12)	0.0 (0)	44.0 (33)	16.0 (12)	24.0 (18)

JFB, % (*n*/ 41 tokens)

	n = 41	≠RL	R	R+L	L
S/N	53.7 (22)	51.2 (21)	0.0 (0)	0.0 (0)	2.4 (1)
M	22.0 (9)	19.5 (8)	2.4 (1)	0.0 (0)	0.0 (0)
W/O	17.1 (7)	--	4.9 (2)	0.0 (0)	12.2 (5)
NOG	7.3 (3)	70.7 (29)	7.3 (3)	0.0 (0)	14.6 (6)

JMB, % (*n*/ 38 tokens)

	n = 38	≠RL	R	R+L	L
S/N	36.8 (14)	7.9 (3)	2.6 (1)	7.9 (3)	18.4 (7)
M	15.8 (6)	2.6 (1)	5.3 (2)	2.6 (1)	5.3 (2)
W/O	36.8 (14)	--	23.7 (9)	0.0 (0)	13.2 (5)
NOG	10.5 (4)	10.5 (4)	31.6 (12)	10.5 (4)	36.8 (14)

The results shown in the tables suggest that, for all speakers other than JMB, there was an overall tendency toward open strictures particularly of the rhotic variety. The rhotic approximant [ɹ] and lowered non-lateral flap [ɾ̞] respectively counted for 38 (21.2%) and 15 (8.4%) of the 179 tokens, a combined total of 29.6% of the transcriptions. Laterals of all stricture types made up 22.9% (*n*= 41) followed closely at 22.3% (*n*= 40) by realizations that were neither lateral nor rhotic. Interestingly, 17 of the 179 tokens (9.5%) were categorized as both lateral and rhotic. These sounds, such as in the pre-rhoticized lateral flap [˞ɺ], exhibited a rhotic quality going into the stricture and a lateral quality going out of it. No instances were observed in which this ordering of rhoticity and laterality was reversed.

As for the speakers individually, 12 of JFA’s 25 tokens were open rhotics: 6 (25%) each of the rhotic approximant [ɹ] and lowered non-lateral flap [ɾ̞]. JFA produced an additional mid-range stricture type rhotic ([˞ɾʴ], not shown in Table 11), bringing her total proportion of rhotics to 52% (*n* = 13). Non-lateral non-rhotics were her next most prevalent, with 3 tokens each of the raised and plain non-lateral flaps [ɾ̝, ɾ] (12% each). The lowered/open lateral flap [ɺ̞] occurred twice (8%).

JFB’s 41 realizations of /r/ were characteristically of the narrow-stricture non-lateral non-rhotic type; 10 (24.4%) were transcribed as the raised non-lateral flap [ɾ̝] while a further 9 (22%) were tapped fricatives [ɾ̝̊], also categorized as strong/narrow non-lateral non-rhotics. The somewhat greater proportion of these transcriptions may be due to tokens where the /ore/ sequence was in the demonstrative pronoun *sore* (‘that’). In these cases, it is likely that frication noise from the initial [s] ‘spread’ through the first vowel to /r/, resulting in the auditory impression of a raised flap. In any case, the combined proportion of this category was 51.2% (*n* = 21; includes two instances of [d] not shown in Table 11). Her next most frequent realizations were the plain non-lateral flap [ɾ] (19.5%, *n*= 8) followed by the lowered/open lateral flap [ɺ̞] (12.2%, *n*= 5). Including one other instance of the short lateral approximant [l̆] (not shown in Table 11), laterals accounted for 14.6% (*n*= 6) of her productions.

As mentioned earlier, JMA's use of the first-person pronoun *ore* likely contributed to his conspicuously large number of tokens, 75. Of these, the single most frequent realization was the rhotic approximant [ɹ] at 33.3% (*n*= 25). Combined with 6 (8%) other instances of the lowered non-lateral flap [ɾ̞] as well as one other narrow-stricture rhotic ([ʵɾ̞], not shown in Table 11), rhotics comprised 44% (*n*= 33) of his realizations. In 12 instances (16%), no oral gesture was perceived. Laterals of various stricture types accounted for 24% (*n*= 18) of JMA's tokens, of which the short lateral approximant [l̆] occurred 5 times (6.7%). Realizations categorized as both rhotic and lateral made of 16% of his productions, with 7 (9.3%) pre-rhoticized (mid-range stricture) lateral flaps [ʵɺ], and 5 pre-rhoticized short lateral approximants [ʵl̆].

Speaker JMB, who generally used the first-person pronoun *boku* to refer to himself, had a total of 38 tokens of /r/ in this environment. Laterals, followed by rhotics, formed the largest proportions of JMB's realizations at 36.8% and 31.6% (*n*= 14 and 12) respectively. Within these categories each of the rhotic and short lateral approximants [ɹ, l̆] occurred 6 times (15.8%), with the lowered/open lateral flap [ɺ̞] occurring 5 times at 13.2%. There were 4 instances (10.5%) where no oral gesture was detected. An even number of realizations (n= 4, 10.5%) of both the non-lateral/non-rhotic and rhotic plus lateral categories also occurred.

4.6.3 /a_i/

By itself, the sequence /ari/ can mean 'ant' or, as the connective *renyoukei* form of the verb *aru,* 'to be' for inanimate objects. Incidentally, the *renyoukei* is the form of verbs to which the polite auxiliary *-masu* attaches (ex. *ari-masu* 'be-POLITE:NONPAST'). This environment is also found in the affix *tari* ('et cetera' for verbial expressions). In total, 166 instances of the /a_i/ environment occurred in the two conversations, with the two females' proportions of tokens being somewhat higher than the males' possibly due to their more frequent use of polite forms. The transcriptions of the 166 tokens are summarized in Table 13, with their categories are summarized in Table 14.

Table 13. Transcriptions for /r/ in /a_i/ (NOG: no audible oral gesture). Values (*n*, %) apply additively to each symbol where more than is listed.

All speakers (*n*= 166; 100%)	JFA (*n*= 49; 100%)	JFB (*n*= 68; 100%)	JMA (*n*= 20; 100%)	JMB (*n*= 29; 100%)
ɾ̞ (*n*= 40; 24.1%)	ɾ̞, ɾ (*n/2*= 13; 26.5%)	ɾ̞ (*n*= 24; 35.3%)	l̆ (*n*= 12; 60.0%)	l̆ (*n*= 8; 27.6%)
ɹ, ɹ̝ (*n/2*= 26; 15.7%)	ɹ̝ (*n*= 10; 20.4%)	ɹ (*n*= 15; 22.1%)	ɹ̝ (*n*= 5; 25.0%)	ɹ (*n*= 4; 13.8%)
ɾ (*n*= 25; 15.1%)	ɹ (*n*= 7; 14.3%)	ɾ (*n*= 10; 14.7%)	ɾ̞, ɾ, l (*n/3*= 1; 5.0%)	ɹ̝, ɾ̞, NOG (*n/3*= 3; 10.3%)
l̆ (*n*= 22; 13.3%)	ɾ̞ (*n*= 3; 6.1%)	ɹ̝ (*n*= 8; 11.8%)		ɾ̞, ʳl̆, NOG (*n/3*= 2; 6.9%)
others (*n*= 27; 16.3%)	others (*n*= 3; 6.1%)	others (*n*= 11; 16.2%)	others (*n*= 5; 16.1%)	others (*n*= 5; 17.2%)

Table 14. Category frequencies for /a_i/ (S/N: strong/narrow, M: mid-range, W/O: weak/open; ≠RL: non-rhotic non-lateral, R: rhotic, R+L: rhotic plus lateral, L: lateral; NOG: no audible oral gesture)

JFA, % (*n*/ 49 tokens)

		≠RL	R	R+L	L
S/N	30.6 (15)	26.5 (13)	0.0 (0)	0.0 (0)	4.1 (2)
M	40.8 (20)	26.5 (13)	0.0 (0)	0.0 (0)	14.3 (7)
W/O	26.5 (13)	--	6.1 (3)	0.0 (0)	20.4 (10)
NOG	2.0 (1)	53.1 (26)	6.1 (3)	0.0 (0)	38.8 (19)
n=	49				

JMA, % (*n*/ 20 tokens)

		≠RL	R	R+L	L
S/N	70.0 (14)	5.0 (1)	0.0 (0)	0.0 (0)	65.0 (13)
M	5.0 (1)	5.0 (1)	0.0 (0)	0.0 (0)	0.0 (0)
W/O	25.0 (5)	--	0.0 (0)	0.0 (0)	25.0 (5)
NOG	0.0 (0)	10.0 (2)	0.0 (0)	0.0 (0)	90.0 (18)
n=	20				

JFB, % (*n*/ 68 tokens)

		≠RL	R	R+L	L
S/N	50.0 (34)	48.5 (33)	0.0 (0)	0.0 (0)	1.5 (1)
M	36.8 (25)	14.7 (10)	0.0 (0)	0.0 (0)	22.1 (15)
W/O	11.8 (8)	--	0.0 (0)	0.0 (0)	11.8 (8)
NOG	1.5 (1)	63.2 (43)	0.0 (0)	0.0 (0)	35.3 (24)
n=	68				

JMB, % (*n*/ 29 tokens)

		≠RL	R	R+L	L
S/N	44.8 (13)	10.3 (3)	0.0 (0)	6.9 (2)	27.6 (8)
M	20.7 (6)	3.4 (1)	0.0 (0)	0.0 (0)	17.2 (5)
W/O	24.1 (7)	--	13.8 (4)	0.0 (0)	10.3 (3)
NOG	10.3 (3)	13.8 (4)	13.8 (4)	6.9 (2)	55.2 (16)
n=	29				

The proportionately most frequent categories in this environment were laterals at 46.4% (*n* = 77) and non-lateral non-rhotics at 45.2% (*n* = 75). Rhotics accounted for only 4.2% of the data, with 7 tokens. Turning to stricture types, there was a general preference for narrow/strong strictures (45.8%, *n* = 76) over mid-range (31.3%, *n* = 52) or open/weak ones (19.9%, n = 33). All speakers did, however, produce higher than average proportions of mid-range and lowered lateral flaps. In terms of individual transcriptions, the raised/narrow non-lateral non-rhotic flap [ɾ̝] was the most frequent among JFB's productions at 35.3% (n = 24), and along with the mid-range non-lateral flap [ɾ] it was one of JFA's most frequent as well at 26.5% (n = 13). Both male speakers' most prevalent realization was the (narrow/strong) short lateral approximant [l̆], which comprised 60% (n = 12) of JMA's and 27.6% (n = 8) of JMB's productions.

4.6.4 /e_ɯ/

Like the other environments examined in so far, /eɾɯ/ features in a range of non-functional lexical items (e.g. *tab<u>e-ru</u>* 'eat-NONPAST') as well as functional/grammatical constructions. Of the latter type, one of the most common is in the casual-speech form of progressive verbial constructions. Progressive constructions in Japanese are formed by affixing the particle *te* and the auxiliary *iru* ('to be' for animate objects) to the connective *renyoukei* form of verbs. For instance the non-past form of 'eat' is *tabe-ru*, and its progressive form is *tabe-te-iru*. In casual speech, more often than not, the sequence *ei* is not realized as a diphthong but rather as /e/. For this reason, progressive constructions in this study's data that did not use the diphthong were treated as instances of /e/. One hundred thirty-two tokens of /r/ occurred in the environment /e_ɯ/, the transcriptions for which are summarized in Table 15 with their categories summarized further below in Table 16.

Table 15. Transcriptions for /r/ in /e_ɯ/ (NOG: no audible oral gesture). Values (*n*, %) apply additively to each symbol where more than is listed.

All speakers (*n*= 132; 100%)	JFA (*n*= 24; 100%)	JFB (*n*= 32; 100%)	JMA (*n*= 41; 100%)	JMB (*n*= 35; 100%)
ɾ (*n*= 28; 21.2%)	ɾ (*n*= 7; 29.2%)	ɾ̝ (*n*= 14; 43.8%)	ɺ (*n*= 9; 22.0%)	ɹ (*n*= 8; 22.9%)
ɾ̝ (*n*= 27; 20.5%)	ɾ̝ (*n*= 6; 25.0%)	ɾ (*n*= 9; 28.1%)	ɾ̞ (*n*= 8; 19.5%)	ɾ (*n*= 6; 17.1%)
ɺ, ɾ̞ (*n/2*= 16; 12.1%)	ɺ (*n*= 4; 16.7%)	ɾ̥ (*n*= 3; 9.4%)	ɾ (*n*= 6; 14.6%)	l̆, ɺ, ɾ̞ (*n/3*= 3; 8.6%)
ɹ (*n*= 11; 8.3%)	ɾ̞ (*n*= 3; 12.5%)	ɾ̞ (*n*= 2; 6.2%)	l, NOG (*n/2*= 4; 9.8%)	ɺ̞ (*n*= 2; 5.7%)
	l (*n*= 2; 8.3%)		l̆ (*n*= 3; 7.3%)	
others (*n*= 34; 25.8%)	others (*n*= 2; 8.3%)	others (*n*= 4; 12.5%)	others (*n*= 7; 17.1%)	others (*n*= 5; 14.3%)

Category-wise, non-lateral non-rhotics with mid-range to narrow stricture types (i.e. [ɾ, ɾ̝, ɾ̥]) made up 62 of the 132 tokens, or 47%. Laterals ([l, l̆, ɺ, ɺ̞]) were the second largest group, with 34 tokens (25.8%). Rhotics ([ɹ, ɾ̞]) were nearly equal in number at 31, or 23.5%.

Looking at each speaker individually, JFA realized /r/ in this environment primarily as mid- to narrow non-lateral flaps: [ɾ, ɾ̝] at 29.2% and 25% (*n*= 7, 6) respectively, 54.2% in total. JFA realized 7 of her 24 tokens (29.2%) as laterals, the largest group of which being the lateral flap [ɺ] (*n* = 4, or 16.7%).

Table 16. Category frequencies for /e_ɯ/ (S/N: strong/narrow, M: mid-range, W/O: weak/open; ≠RL: non-rhotic non-lateral, R: rhotic, R+L: rhotic plus lateral, L: lateral; NOG: no audible oral gesture)

JFA, % (*n*/ 24 tokens)

S/N	33.3 (8)	25.0 (6)	0.0 (0)	0.0 (0)	8.3 (2)
M	45.8 (11)	29.2 (7)	0.0 (0)	0.0 (0)	16.7 (4)
W/O	20.8 (5)	--	16.7 (4)	0.0 (0)	4.2 (1)
NOG	0.0 (0)	54.2 (13)	16.7 (4)	0.0 (0)	29.2 (7)
n =	24	≠RL	R	R+L	L

JMA, % (*n*/ 41 tokens)

S/N	24.4 (10)	7.3 (3)	0.0 (0)	0.0 (0)	17.1 (7)
M	39.0 (16)	14.6 (6)	2.4 (1)	0.0 (0)	22.0 (9)
W/O	26.8 (11)	--	24.4 (10)	0.0 (0)	2.4 (1)
NOG	9.8 (4)	22.0 (9)	26.8 (11)	0.0 (0)	41.5 (17)
n =	41	≠RL	R	R+L	L

JFB, % (*n*/ 32 tokens)

S/N	59.4 (19)	59.4 (19)	0.0 (0)	0.0 (0)	0.0 (0)
M	31.2 (10)	28.1 (9)	3.1 (1)	0.0 (0)	0.0 (0)
W/O	9.4 (3)	--	6.2 (2)	0.0 (0)	3.1 (1)
NOG	0.0 (0)	87.5 (28)	9.4 (3)	0.0 (0)	3.1 (1)
n =	32	≠RL	R	R+L	L

JMB, % (*n*/ 35 tokens)

S/N	25.7 (9)	17.1 (6)	0.0 (0)	0.0 (0)	8.6 (3)
M	31.4 (11)	17.1 (6)	5.7 (2)	0.0 (0)	8.6 (3)
W/O	42.9 (15)	--	34.1 (11)	2.9 (1)	8.6 (3)
NOG	0.0 (0)	34.3 (12)	37.1 (13)	2.9 (1)	25.7 (9)
n =	35	≠RL	R	R+L	L

JFB remained true to form in that the greatest proportion (42%, *n*= 7) of her 32 /r/s in this environment were realized as raised non-lateral flaps; a further 3 of her tokens were the similarly categorized tapped fricative [ɾ̝̊]. Combined, these two transcriptions account for 53.1% (*n*= 17) of her tokens. Further combining JFB's second most frequent transcription type, mid-range non-lateral flaps (28.1%, *n*= 9), the non-lateral non-rhotic category amounts to 87.5 (*n*= 28) of her tokens.

Quite different from the two female speakers, JMA showed an overall preference for lateral realizations for his 41 /r/s. Laterals of all stricture types totalled 41.5% (*n*= 17) of his productions, the largest type of which was the lateral flap [ɺ] at 22% (*n*= 9). JMA's second most frequent realization was the lowered non-lateral flap [ɾ̞], an open rhotic, at 19.5% (*n*= 8). JMA produced a further 3 rhotic tokens, bringing the combined proportion of this category to 26.8% of his tokens. Non-lateral non-rhotics, his most frequent being the plain flap [ɾ] at14.6% (*n*= 6), accounted for 22% of JMA's realizations at a total of 9 tokens.

JMB's 35 realizations were divided nearly evenly among rhotics (13 at 37.1%), non-rhotic non-laterals (12 at 34.3%), and to a slightly lesser extent laterals (9 at 25.7%). The single most frequently transcribed realization was the rhotic approximant [ɹ] at 22.9% (*n*= 8) followed by the mid-range non-lateral (or 'plain') flap [ɾ] at 17.1% (*n*= 6). The lowered, or open non-lateral flap [ɾ̞] (categorized as a rhotic) occurred 3 times (8.6%), as did each of the short lateral approximant [l̆] and the lateral flap [ɺ]. Lastly, the lowered (i.e. open) lateral flap occurred twice, making up 5.7% of JMB's data in this environment.

4.6.5 /a_ɯ/

On its own, the sequence /arɯ/ is the non-past form of the verb of existence ('to be') used for inanimate objects. Aside from in this functional word, the sequence can also be found in any number of verbs, nouns, and other non-functional lexical items such as *Nukarumiyaki*, the brand name for a particular kind of seafood pancake that is discussed at some length by the two male speakers. The total 117 transcriptions for /r/ in /a_ɯ/ in the present dataset are summarized overleaf in Table 17, with their corresponding categories given further below in Table 18.

Table 17. Transcriptions for /r/ in /a_ɯ/ (NOG: no audible oral gesture). Values (*n*, %) apply additively to each symbol where more than is listed.

All speakers (*n*= 117; 100%)	JFA (*n*= 21; 100%)	JFB (*n*= 21; 100%)	JMA (*n*= 37; 100%)	JMB (*n*= 38; 100%)
ɾ̞ (*n*= 18; 15.4%)	ɾ̞ (*n*= 7; 33.3%)	ɾ̞ (*n*= 9; 42.9%)	ɾ̞ (*n*= 9; 24.3%)	l̆, ɹ (*n/2*= 7; 18.4%)
ɹ (*n*= 17; 14.5%)	ɾ̞, ɹ (*n/2*= 3; 14.3%)	ɾ (*n*= 4; 19.0%)	ɺ, ɹ (*n/2*= 7; 18.9%)	ɺ (*n*= 6; 15.8%)
ɺ (*n*= 14; 12.0%)	ɾ (*n*= 2; 9.5%)	ɾ̥ (*n*= 3; 14.3%)	l, l̆, ˞ɾ, ɾʴ (*n/4*= 3; 8.1%)	l, ɾ̞, ˞ɾ (*n/3*= 2; 5.3%)
ɾ̞ (*n*= 13; 11.1%)				
l̆ (*n*= 11; 9.4%)				
ɾ (*n*= 7; 6.0%)				
ɾʴ, l (*n/2*= 6; 5.1%)				
others (*n*= 25; 21.4%)	others (*n*= 6; 28.6%)	others (*n*= 5; 23.8%)	others (*n*= 2; 5.4%)	others (*n*= 12; 31.6%)

The most common transcriptions for /r/ were the lowered non-lateral flap [ɾ̞] at 15.4% (*n*= 18), followed by the rhotic approximant [ɹ] at 14.5% (*n*= 17). As discussed earlier in the description of the perceptual categories used in this study, both transcriptions are classified as *open rhotics* in that their auditory impression is similar to that associated with a rising or falling F3 in North American English (the author's first language and perceptual reference). These transcriptions rank first and second most frequent for speakers JFA and JMA, while the rhotic approximant [ɹ] shares JMB's 'most frequent' spot with the short lateral approximant [l̆]. Speaker JFB's productions differ from the other three speakers in that the great majority of her /r/s are narrow non-lateral non-rhotics, most frequent of which being the raised non-lateral flap [ɾ̝] (42.9%, *n*= 9). Further consideration is given to category frequencies in Table 18 below.

Table 18. Category frequencies for /a_ɯ/ (S/N: strong/narrow, M: mid-range, W/O: weak/open; ≠RL: non-rhotic non-lateral, R: rhotic, R+L: rhotic plus lateral, L: lateral; NOG: no audible oral gesture)

JFA, % (*n*/ 21 tokens)

S/N	28.6 (6)	23.8 (5)	0.0 (0)	0.0 (0)	4.8 (1)
M	23.8 (5)	9.5 (2)	9.5 (2)	4.8 (1)	0.0 (0)
W/O	47.6 (10)	--	47.6 (10)	0.0 (0)	0.0 (0)
NOG	0.0 (0)	33.3 (7)	57.1 (12)	4.8 (1)	4.8 (1)
n=	21	**≠RL**	**R**	**R+L**	**L**

JMA, % (*n*/ 37 tokens)

S/N	16.2 (6)	0.0 (0)	0.0 (0)	0.0 (0)	16.2 (6)
M	35.1 (13)	0.0 (0)	16.2 (6)	0.0 (0)	18.9 7)
W/O	45.9 (17)	--	43.2 (16)	0.0 (0)	2.7 (1)
NOG	2.7 (1)	0.0 (0)	59.5 (22)	0.0 (0)	37.8 (14)
n=	37	**≠RL**	**R**	**R+L**	**L**

JFB, % (*n*/ 21 tokens)

S/N	66.7 (14)	61.9 (13)	0.0 (0)	0.0 (0)	4.8 (1)
M	28.6 (6)	19.0 (4)	4.8 (1)	0.0 (0)	4.8 (1)
W/O	4.8 (1)	--	0.0 (0)	0.0 (0)	4.8 (1)
NOG	0.0 (0)	81.0 (17)	4.8 (1)	0.0 (0)	14.3 (3)
n=	21	**≠RL**	**R**	**R+L**	**L**

JMB, % (*n*/ 38 tokens)

S/N	34.2 (13)	5.3 (2)	2.6 (1)	0.0 (0)	26.3 (10)
M	28.9 (11)	2.6 (1)	7.9 (3)	2.6 (1)	15.8 (6)
W/O	36.8 (14)	--	28.9 (11)	0.0 (0)	7.9 (3)
NOG	0.0 (0)	7.9 (3)	39.5 (15)	2.6 (1)	50.0 (19)
n=	38	**≠RL**	**R**	**R+L**	**L**

Considered together with other transcriptions classified as rhotic, such as pre- and post-rhoticized non-lateral flaps ([ʴɾ, ɾʴ]), rhotics account for 42.7% (*n*= 50) of all speakers' tokens in this environment. While less frequent, especially for the female speakers, laterals make up 31.6% (n= 37) of the tokens. In terms of the central oral stricture, productions with narrow/strong strictures make up a third (33.3%, n= 39) of the tokens while open/weak strictures account for 35.9% (n= 42) of the data, with the remainder (29.9%, n= 35) classified as mid-range.

4.6.6 /a_e/

A total of 107 tokens of /r/ occurred in this environment, and the transcriptions of these are summarized in Table 19 with their corresponding categories further below in Table 20.

Table 19. Transcriptions for /r/ in /are/ (NOG: no audible oral gesture). Values (*n*, %) apply additively to each symbol where more than is listed.

All speakers (*n*= 107; 100%)	JFA (*n*= 30; 100%)	JFB (*n*= 24; 100%)	JMA (*n*= 31; 100%)	JMB (*n*= 22; 100%)
ɺ̞ (*n*= 22; 20.6%)	ɺ̞ (*n*= 15; 50.0%)	ɾ̞ (*n*= 11; 45.8%)	NOG (*n*= 9; 29.0%)	ɺ̆ (*n*= 5; 22.7%)
ɺ̆ (*n*= 16; 15.0%)	ɾ̞ (*n*= 5; 16.7%)	ɺ̆, ɾ (*n/2*= 3; 12.5%)	ɺ̆ (*n*= 7; 22.6%)	ɹ (*n*= 4; 18.2%)
ɾ̞, ɺ (*n/2*= 12; 11.2%)	ɺ (*n*= 4; 13.3%)	ɾ̥, ɺ, ɺ̞ (*n/3*= 2; 8.3%)	ɺ (*n*= 5; 16.1%)	ɾ̞ (*n*= 3; 13.6%)
NOG (*n*= 11; 10.3%)	ɾ (*n*= 3; 10.0%)		ɺ̞ (*n*= 3; 9.7%)	ɺ̞, ˜ɺ, NOG (*n/3*= 2; 9.1%)
ɾ̞ (*n*= 9; 8.4%)			ɹ (*n*= 2; 6.5%)	
ɾ (*n*= 7; 6.5%)				
ɹ (*n*= 6; 5.6%)				
others (*n*= 12; 11.2%)	others (*n*= 3; 10.0%)	others (*n*= 1; 4.2%)	others (*n*= 5; 16.1%)	others (*n*= 4; 18.2%)

Aside from appearing in any number of lexical contexts including the interjection *are* ('What the..?'), the /are/ sequence can also be found in the passive forms of verbs: ex. *su-ru* ('do-NONPAST') versus *sare-ru* ('do:PSV-NONPAST').

Table 20. Category frequencies for /a_e/ (S/N: strong/narrow, M: mid-range, W/O: weak/open; ≠RL: non-rhotic non-lateral, R: rhotic, R+L: rhotic plus lateral, L: lateral; NOG: no audible oral gesture)

JFA, % (*n*/ 30 tokens)

S/N	10.0 (3)	6.7 (2)	0.0 (0)	0.0 (0)	3.3 (1)
M	23.3 (7)	10.0 (3)	0.0 (0)	0.0 (0)	13.3 (4)
W/O	66.7 (20)	--	16.7 (5)	0.0 (0)	50.0 (15)
NOG	0.0 (0)	16.7 (5)	16.7 (5)	0.0 (0)	66.7 (20)
n =	30	≠RL	R	R+L	L

JMA, % (*n*/ 31 tokens)

S/N	32.3 (10)	0.0 (0)	0.0 (0)	6.5 (2)	25.8 (8)
M	22.6 (7)	3.2 (1)	0.0 (0)	3.2 (1)	16.1 (5)
W/O	16.1 (5)	--	6.5 (2)	0.0 (0)	9.7 (3)
NOG	29.0 (9)	3.2 (1)	6.5 (2)	9.7 (3)	51.6 (16)
n =	31	≠RL	R	R+L	L

JFB, % (*n*/ 24 tokens)

S/N	66.7 (16)	54.2 (13)	0.0 (0)	0.0 (0)	12.5 (3)
M	20.8 (5)	12.5 (3)	0.0 (0)	0.0 (0)	8.3 (2)
W/O	12.5 (3)	--	4.2 (1)	0.0 (0)	8.3 (2)
NOG	0.0 (0)	66.7 (16)	4.2 (1)	0.0 (0)	29.2 (7)
n =	24	≠RL	R	R+L	L

JMB, % (*n*/ 22 tokens)

S/N	31.8 (7)	0.0 (0)	0.0 (0)	4.5 (1)	27.3 (6)
M	13.6 (3)	0.0 (0)	0.0 (0)	9.1 (2)	4.5 (1)
W/O	45.5 (10)	--	36.4 (8)	0.0 (0)	9.1 (2)
NOG	9.1 (2)	0.0 (0)	36.4 (8)	13.6 (3)	40.9 (9)
n =	22	≠RL	R	R+L	L

The overall picture that the results in the tables presents is somewhat mixed; however, there does seem to be a general preference for lateral realizations over rhotics and non-rhotic/non-laterals – which in turn supports Akamatsu's (1997) view that the pre-/e/ environment is amenable to laterality. Leaving the nature of central strictures aside for the moment, laterals accounted for 48.6% (*n* = 52) of the tokens in /are/ sequences. If we combine these laterals with realizations which were simultaneously lateral as well as rhotic (a further 6 tokens), this figure jumps to 54.2%. Of course, laterals were not the only realizations produced here. Again setting stricture types aside, non-lateral non-rhotic tokens accounted for 20.6% (*n* = 22) of the 107 while rhotics made up 15% (*n* = 16) of the data (20.6% including the 6 cross-category tokens).

In terms of each speakers' individual patterns, half (*n* = 15, 50%) of JFA's realizations were transcribed as the lowered (weak/open) lateral flap followed by the lowered (weak/open) non-lateral flap (16.7%). The mid-range stricture type analogues of these two transcriptions ([ɺ, ɾ]) constituted 13.3% and 10% of JFA's tokens respectively. For speaker JFA, then, this suggests that she tended toward lateral above non-lateral flaps, and of these she tended towards open central oral strictures over mid-range types.

Speaker JFB’s realizations characteristically tended towards the raised (i.e. narrow/strong stricture) non-lateral flap [ɾ̝] followed by its mid-range analogue [ɾ] at 45.8% (*n*= 11) and 12.5% (*n*= 3) respectively. A further 2 (8.3%) of JFB’s realizations were transcribed as tapped fricatives ([ɾ̝̊]). Combined, these add up to 54.2% (*n*= 13) of JFB’s tokens. While JFB clearly preferred narrow non-lateral/non-rhotic realizations, she did produce some laterals nonetheless. Stricture types aside, laterals (7 tokens) accounted for 29.2% of her productions. Her results in Table 20 also suggest higher than average proportions for laterals of all stricture types, as well as a lower than average proportion of strong/narrow non-rhotic non-laterals.

As for the male speakers, JMA most frequently produced no oral gesture at all (29%, *n*= 9). His next most frequent were laterals of sequentially more open stricture categories: [l̆, ɺ, ɺ̞] at 22.6%, 16.1%, and 9.5% (*n*= 7, 5, 3) respectively. In all, laterals accounted for 51.6% (*n*= 16) of his productions while rhotics made up 6.5% (*n*= 2); a further 3 tokens (9.7%) were classified as simultaneously rhotic and lateral. Of speaker JMB’s 22 tokens, the greater proportion were laterals, with the short lateral approximant (with narrow/strong central stricture) being most frequent at 5 tokens (22.7%). Stricture types aside, however, JMB produced laterals (*n*= 9, 40.9%) and rhotics (*n*= 8, 36.4%) in nearly even amounts.

4.7 Specific vocalic environments

The Sections that follow present transcription and category both before and following each of the five vowels of Japanese: /i, ɯ, e, o, a/.

4.7.1 /_i/

A total of 255 tokens from the 1,535 examined here occurred before the high front vowel. This number includes instances of underlying consonant-glide sequences such as in the word *ryuu* ('flow') and *ryo* ('both'). Recall that Akamatsu (1997, 1999) considers /r/ in these circumstances to be distinct palatalized variants of whichever tap or lateral a speaker may produce in other contexts. The decision to group these into the /_i/ environment is made here based on articulatory considerations: as [j] is much more similar to [i] than it is to either [o] or [ɯ], its inclusion in the /_i/ environment seemed a more accurate depiction of the gestures surrounding the segment of foremost interest, /r/. Table 21 summarizes the transcriptions of each speaker's realizations of /r/ before [i] which accounted for more than 5% of his or her productions; category results are summarized in Table 22.

Table 21. Transcriptions for /r/ before /i/ (NOG: no audible oral gesture). Values (*n*, %) apply additively to each symbol where more than is listed.

All speakers (*n*= 255; 100%)	JFA (*n*= 67; 100%)	JFB (*n*= 90; 100%)	JMA (*n*= 48; 100%)	JMB (*n*= 50; 100%)
ɾ̝ (*n*= 63; 24.7%)	ɾ̝ (*n*= 17; 25.4%)	ɾ̝ (*n*= 37; 41.1%)	ɺ̆ (*n*= 15; 31.3%)	ɺ̆ (*n*= 11; 22.0%)
ɺ̞ (*n*= 38; 14.9%)	ɾ (*n*= 16; 23.9%)	ɺ (*n*= 16; 17.8%)	ɺ̞ (*n*= 14; 29.2%)	ɾ̝ (*n*= 8; 16.0%)
ɺ (*n*= 34; 13.3%)	ɺ̞ (*n*= 12; 17.9%)	ɾ (*n*= 14; 15.6%)	ɺ (*n*= 4; 8.3%)	ɺ (*n*= 7; 14.0%)
ɾ (*n*= 33; 12.9%)	ɾ̞ (*n*= 8; 11.9%)	ɺ̞ (*n*= 8; 8.9%)	ɾ̞ (*n*= 3; 6.3%)	NOG (*n*= 5; 10.0%)
ɺ̆ (*n*= 29; 11.4%)	ɺ (*n*= 7; 10.4%)	ɾ̥ (*n*= 5; 5.6%)		ɺ̞, ɾ̞ (*n/2*= 4; 8.0%)
ɾ̞ (*n*= 15; 5.9%)	NOG (*n*= 4; 6.0%)			
others (*n*= 43; 16.9%)	others (*n*= 7; 10.4%)	others (*n*= 10; 11.1%)	others (*n*= 11; 22.9%)	others (*n*= 11; 22.0%)

With the exception of speaker JFB, the majority of whose tokens strong/narrow raised flaps, all of the speakers produced a wide range of /r/ variants before [i]. No speaker produced any single variant over fifty percent of the time. Common to all speakers' most-frequent variants, however, is the notion of a strong/narrow central oral closure – the difference being that the two genders separately opted for either lateral or non-rhotic non-lateral accompaniments to that oral closure.

Table 22. Category frequencies before /i/ (S/N: strong/narrow, M: mid-range, W/O: weak/open; ≠RL: non-rhotic non-lateral, R: rhotic, R+L: rhotic plus lateral, L: lateral; NOG: no audible oral gesture)

JFA, % (*n*/ 67 tokens)

	n= 67	≠RL	R	R+L	L
S/N	29.9 (20)	26.9 (18)	0.0 (0)	0.0 (0)	3.0 (2)
M	34.3 (23)	23.9 (16)	0.0 (0)	0.0 (0)	10.4 (7)
W/O	29.9 (20)	--	11.9 (8)	0.0 (0)	17.9 (12)
NOG	6.0 (4)	50.7 (34)	11.9 (8)	0.0 (0)	31.3 (21)

JMA, % (*n*/ 48 tokens)

	n= 48	≠RL	R	R+L	L
S/N	43.8 (21)	4.2 (2)	0.0 (0)	0.0 (0)	39.6 (19)
M	12.5 (6)	2.1 (1)	2.1 (1)	0.0 (0)	8.3 (4)
W/O	37.5 (18)	--	8.3 (4)	0.0 (0)	29.2 (14)
NOG	6.3 (3)	6.3 (3)	10.4 (5)	0.0 (0)	77.1 (37)

JFB, % (*n*/ 90 tokens)

	n= 90	≠RL	R	R+L	L
S/N	56.7 (51)	54.4 (49)	0.0 (0)	0.0 (0)	2.2 (2)
M	33.3 (30)	15.6 (14)	0.0 (0)	0.0 (0)	17.8 (16)
W/O	8.9 (8)	--	0.0 (0)	0.0 (0)	8.9 (8)
NOG	1.1 (1)	70.0 (63)	0.0 (0)	0.0 (0)	28.9 (26)

JMB, % (*n*/ 50 tokens)

	n= 50	≠RL	R	R+L	L
S/N	48.0 (24)	20.0 (10)	2.0 (1)	4.0 (2)	22.0 (11)
M	24.0 (12)	4.0 (2)	0.0 (0)	4.0 (2)	16.0 (8)
W/O	18.0 (9)	--	10.0 (5)	0.0 (0)	8.0 (4)
NOG	10.0 (5)	24.0 (12)	12.0 (6)	8.0 (4)	46.0 (23)

Comparing the category frequency results in Table 22 with each speaker's averages for the entire dataset, an additional observation can be made. All four speakers produced fewer rhotics before [i]: 11.9% (*n*= 8) of JFA's 67 tokens were rhotic, compared to her 30.5% for the same category for all environments. JMA produced 10.4% (*n*= 5) rhotics before [i] compared to his 37.6% average across all environments. Similarly, JMB's proportion of rhotics before [i] fell to 12.0% (*n*= 6) from his overall overage of 40.7%.

JFB produced no rhotics before [i] as opposed to 6% of her productions across all environments.

4.7.2 /i_/

Similar to how sequences like [jo] and [ja] were included in the /_i/ environment due to the articulatory similarity between [j] and [i], a similar inclusion was made of pre-/r/ diphthongs (/eʲ, aʲ, oʲ, ɯʲ/) into the /i_/ environment. Transcriptions for the 159 post-[i] tokens are summarized in Table 23, and their perceptual categories are given in further below in Table 24.

Table 23. Transcriptions for /r/ after /i/ (NOG: no audible oral gesture). Values (*n*, %) apply additively to each symbol where more than is listed.

All speakers **(*n*= 159; 100%)**	**JFA** **(*n*= 40; 100%)**	**JFB** **(*n*= 34; 100%)**	**JMA** **(*n*= 47; 100%)**	**JMB** **(*n*= 38; 100%)**
ɾ (*n*= 27; 17.0%)	ɾ (*n*= 10; 25.0%)	ɾ̞ (*n*= 13; 38.2%)	ɺ̝ (*n*= 7; 14.9%)	ɹ (*n*= 8; 21.1%)
ɾ̞ (*n*= 19; 11.9%)	ɾ̝ (*n*= 6; 15.0%)	ɾ̥, ɾ (*n/2*= 7; 20.6%)	l̆, ɺ, ɾ̝, NOG (*n/4*= 6; 12.8%)	ɾ (*n*= 5; 13.2%)
ɺ, ɾ̝ (*n/2*= 16; 10.1%)	ɺ (*n*= 5; 12.5%)		ɾ (*n*= 5; 10.6%)	ɺ, NOG (*n/2*= 4; 10.5%)
ɺ̝, ɹ (*n/2*= 14; 8.8%)	ɾ̞, ɾ̥ (*n/2*= 4; 10.0%)		ɾʴ, ɹ (*n/2*= 3; 6.3%)	l̆, ɺ̝, ɾ̝ (*n/3*= 3; 7.9%)
ɾ̥ (*n*= 13; 8.2%)	ɺ̝, ɹ (*n/2*= 3; 7.5%)			ɾ̞, ɾ̥ (*n/2*= 2; 5.3%)
NOG (*n*= 11; 6.9%)	ɾ̞ʴ (*n*= 2; 5.0%)			
l̆ (*n*= 10; 6.3%)				
others (*n*= 19; 11.9%)	others (*n*= 3; 7.5%)	others (*n*= 7; 20.6%)	others (*n*= 5; 10.6%)	others (*n*= 4; 10.5%)

Compared with their averages for all environments, the four speakers patterned somewhat differently. JFA, JFB, and JMB produced fewer laterals overall in favour of more mid- to narrow/close non-rhotic non-lateral variants as well as narrow rhotics. In contrast, JFB had higher than average proportions of mid-range and strong/narrow non-rhotic non-laterals.

Table 24. Category frequencies following /i/ (S/N: strong/narrow, M: mid-range, W/O: weak/open; ≠RL: non-rhotic non-lateral, R: rhotic, R+L: rhotic plus lateral, L: lateral; NOG: no audible oral gesture)

JFA, % (*n*/ 40 tokens)

n =	40	≠RL	R	R+L	L
S/N	30.0 (12)	25.0 (10)	5.0 (2)	0.0 (0)	0.0 (0)
M	37.5 (15)	25.0 (10)	0.0 (0)	0.0 (0)	12.5 (5)
W/O	30.0 (12)	--	22.5 (9)	0.0 (0)	7.5 (3)
NOG	2.5 (1)	50.0 (20)	27.5 (11)	0.0 (0)	20.0 (8)

JMA, % (*n*/ 47 tokens)

n =	47	≠RL	R	R+L	L
S/N	19.1 (9)	0.0 (0)	0.0 (0)	0.0 (0)	19.1 (9)
M	29.8 (14)	10.6 (5)	6.4 (3)	0.0 (0)	12.8 (6)
W/O	36.2 (17)	--	19.1 (9)	0.0 (0)	17.0 (8)
NOG	14.9 (7)	10.6 (5)	25.5 (12)	0.0 (0)	48.9 (23)

JFB, % (*n*/ 41 tokens)

n =	41	≠RL	R	R+L	L
S/N	67.6 (23)	61.8 (21)	2.9 (1)	0.0 (0)	2.9 (1)
M	26.5 (9)	20.6 (7)	2.9 (1)	0.0 (0)	2.9 (1)
W/O	5.9 (2)	--	2.9 (1)	0.0 (0)	2.9 (1)
NOG	0.0 (0)	82.4 (28)	8.8 (3)	0.0 (0)	8.8 (3)

JMB, % (*n*/ 38 tokens)

n =	38	≠RL	R	R+L	L
S/N	26.3 (10)	15.8 (6)	2.6 (1)	0.0 (0)	7.9 (3)
M	26.3 (10)	13.2 (5)	0.0 (0)	0.0 (0)	13.2 (5)
W/O	36.8 (14)	--	28.9 (11)	0.0 (0)	7.9 (3)
NOG	10.5 (4)	28.9 (11)	31.6 (12)	0.0 (0)	28.9 (11)

4.7.3 /_ɯ/

A total 330 of the total 1,535 tokens examined here occurred before the high-back vowel /ɯ/. Table 25 summarizes the transcriptions of each speaker's realizations of /r/ before /ɯ/ which accounted for more than 5% of his or her productions; corresponding auditory-perceptual categories are summarized further below in Table 26.

Table 25. Transcriptions for /r/ before /ɯ/ (NOG: no audible oral gesture). Values (*n*, %) apply additively to each symbol where more than is listed.

All speakers (*n*= 330; 100%)	JFA (*n*= 56; 100%)	JFB (*n*= 73; 100%)	JMA (*n*= 107; 100%)	JMB (*n*= 94; 100%)
ɾ̝ (*n*= 51; 15.5%)	ɾ̞ (*n*= 14; 25.0%)	ɾ̝ (*n*= 31; 42.5%)	ɺ (*n*= 22; 20.6%)	ɹ (*n*= 18; 19.1%)
ɾ (*n*= 48; 14.5%)	ɾ̝ (*n*= 11; 19.6%)	ɾ (*n*= 16; 21.9%)	ɾ̞ (*n*= 21; 19.6%)	ɾ (*n*= 13; 13.8%)
ɾ̞ (*n*= 46; 13.9%)	ɾ (*n*= 9; 16.1%)	ɾ̥ (*n*= 11; 15.1%)	ɹ (*n*= 12; 11.2%)	ɺ (*n*= 11; 11.7%)
ɺ (*n*= 39; 11.8%)	ɹ (*n*= 5; 8.9%)		ɾ (*n*= 10; 9.3%)	l̆ (*n*= 10; 10.6%)
ɹ (*n*= 35; 10.6%)	ɾ̥, ɺ (*n/2*= 4; 7.1%)		l, NOG (*n/2*= 9; 8.4%)	ɾ̞ (*n*= 9; 9.6%)
l̆ (*n*= 21; 6.4%)	l (*n*= 3; 5.4%)		l̆ (*n*= 7; 6.5%)	ɾ̝ (*n*= 7; 7.4%)
ɾ̥ (*n*= 18; 5.5%)				
others (*n*= 72; 21.8%)	others (*n*= 4; 7.1%)	others (*n*= 15; 20.5%)	others (*n*= 17; 15.9%)	others (*n*= 26; 27.7%)

No single variant in Table 25 accounted for more than half of any speaker's productions, although speaker JFB's use of raised non-lateral flaps was quite substantial at 42.5%. As can be seen in shaded versus underlined cells in Table 26, and compared with their averages across the entire dataset given in Table 7, all of the speakers produced higher percentages of mid- and strong/narrow non-rhotic non-lateral variants.

Table 26. Category frequencies before /ɯ/ (S/N: strong/narrow, M: mid-range, W/O: weak/open; ≠RL: non-rhotic non-lateral, R: rhotic, R+L: rhotic plus lateral, L: lateral; NOG: no audible oral gesture)

	JFA, % (*n*/ 56 tokens)				
S/N	33.9 (19)	26.8 (15)	0.0 (0)	0.0 (0)	7.1 (4)
M	28.6 (16)	16.1 (9)	3.6 (2)	1.8 (1)	7.1 (4)
W/O	37.5 (21)	--	33.9 (19)	0.0 (0)	3.6 (2)
NOG	0.0 (0)	42.9 (24)	37.5 (21)	1.8 (1)	17.9 (10)
n =	56	**≠RL**	**R**	**R+L**	**L**

	JMA, % (*n*/ 107 tokens)				
S/N	19.6 (21)	2.8 (3)	0.0 (0)	0.0 (0)	16.8 (18)
M	38.3 (41)	9.3 (10)	6.5 (7)	1.9 (2)	20.6 (22)
W/O	33.6 (36)	--	30.8 (33)	0.0 (0)	2.8 (3)
NOG	8.4 (9)	12.1 (13)	37.4 (40)	1.9 (2)	40.2 (43)
n =	107	**≠RL**	**R**	**R+L**	**L**

	JFB, % (*n*/ 73 tokens)				
S/N	65.8 (48)	61.6 (45)	0.0 (0)	0.0 (0)	4.1 (3)
M	27.4 (20)	21.9 (16)	2.7 (2)	0.0 (0)	2.7 (2)
W/O	6.8 (5)	--	2.7 (2)	0.0 (0)	4.1 (3)
NOG	0.0 (0)	83.6 (61)	5.5 (4)	0.0 (0)	11.0 (8)
n =	73	**≠RL**	**R**	**R+L**	**L**

	JMB, % (*n*/ 94 tokens)				
S/N	25.5 (24)	9.6 (9)	2.1 (2)	0.0 (0)	13.8 (13)
M	33.0 (31)	13.8 (13)	6.4 (6)	1.1 (1)	11.7 (11)
W/O	39.4 (37)	--	30.9 (29)	1.1 (1)	7.4 (7)
NOG	2.1 (2)	23.4 (22)	39.4 (37)	2.1 (2)	33.0 (31)
n =	94	**≠RL**	**R**	**R+L**	**L**

Also, in spite of three out of the four (all save JFA) producing fewer rhotics overall, all speakers produced higher proportions of mid-rhotics (i.e. non-lateral flaps preceded by rhoticized vowels or followed by rhotic off-glides). In general, all speakers also produced fewer laterals in this environment, although both males used more mid-range laterals (i.e.

lateral flaps) while JFA's proportion of strong/narrow laterals was higher than her average at 7.1% (*n*=4).

4.7.4 /ɯ_/

In contrast to the pre-/ɯ/ environment where the four speakers produced fewer rhotic variants compared to their overall averages for the entire dataset, the post-/ɯ/ environment featured greater proportions of rhotics. Transcription frequencies for each speaker are presented in Table 27, while the perceptual categories which they represent are summarized further below in Table 28.

Table 27. Transcriptions for /ɾ/ following /ɯ/. Values (*n*, %) apply additively to each symbol where more than is listed.

All speakers (*n*= 199; 100%)	JFA (*n*= 25; 100%)	JFB (*n*= 61; 100%)	JMA (*n*= 68; 100%)	JMB (*n*= 46; 100%)
ɹ (*n*= 38; 19.1%)	ɾ̞ (*n*= 10; 40.0%)	ɾ̝ (*n*= 20; 32.8%)	ɹ (*n*= 20; 29.4%)	ɹ (*n*= 16; 34.8%)
ɾ̞ (*n*= 35; 17.6%)	ɾ̝ (*n*= 4; 16.0%)	ɾ̥ (*n*= 12; 19.7%)	ɾ̞ (*n*= 13; 19.1%)	ɾ̞ (*n*= 8; 17.4%)
ɾ̝ (*n*= 31; 15.6%)	ɾ̥ (*n*= 3; 12.0%)	ɾ (*n*= 8; 13.1%)	ɺ (*n*= 10; 14.7%)	ɾ̝ (*n*= 5; 10.9%)
ɾ̥ (*n*= 15; 7.5%)	ɹ (*n*= 2; 8.0%)	l̆ (*n*= 5; 8.2%)	ɺ̞, ɾʴ (*n/2*= 7; 10.3%)	l̆, ɾ, ɾʴ (*n/3*= 3; 6.5%)
ɾ, ɺ (*n/2*= 13; 6.5%)		ɾ̞ (*n*= 4; 6.6%)		
l̆ (*n*= 12; 6.0%)				
ɾʴ, ɺ̞ (*n/2*= 10; 5.0%)				
others (*n*= 22; 11.1%)	others (*n*= 6; 24.0%)	others (*n*= 12; 19.7%)	others (*n*= 11; 16.2%)	others (*n*= 8; 17.4%)

The two male speakers, JMA and JMB, both produced the rhotic approximant [ɹ] most frequently (29.4% and 34.8%; *n*= 20, 16) followed by the lowered non-lateral flap [ɾ̞], also a weak/open rhotic (19.1% and 17.4%; *n*= 14, 8). In contrast, speaker JFB produced 20 instances of raised non-lateral flaps [ɾ̝] (32.8%) followed by 12 (19.7%) tapped fricatives [ɾ̝̊] and 8 (13.1%) mid-range non-lateral flaps [ɾ]. Speaker JFA realized 10 (40.0%) of her 25 tokens as lowered non-lateral flaps, followed by 4 (16.0%) raised non-lateral flaps and 3 (12.0%) tapped fricatives.

Table 28. Category frequencies following /ɯ/ (S/N: strong/narrow, M: mid-range, W/O: weak/open; ≠RL: non-rhotic non-lateral, R: rhotic, R+L: rhotic plus lateral, L: lateral; NOG: no audible oral gesture)

JFA, % (*n*/ 25 tokens)

S/N	44.0 (11)	32.0 (8)	4.0 (1)	0.0 (0)	8.0 (2)
M	4.0 (1)	4.0 (1)	0.0 (0)	0.0 (0)	0.0 (0)
W/O	48.0 (12)	--	48.0 (12)	0.0 (0)	0.0 (0)
NOG	4.0 (1)	36.0 (9)	52.0 (13)	0.0 (0)	8.0 (2)
n=	25	**≠RL**	**R**	**R+L**	**L**

JMA, % (*n*/ 68 tokens)

S/N	11.8 (8)	2.9 (2)	0.0 (0)	0.0 (0)	8.8 (6)
M	27.9 (19)	1.5 (1)	11.8 (8)	0.0 (0)	14.7 (10)
W/O	58.8 (40)	--	48.6 (33)	0.0 (0)	10.3 (7)
NOG	1.5 (1)	4.4 (3)	60.3 (41)	0.0 (0)	33.8 (23)
n=	68	**≠RL**	**R**	**R+L**	**L**

JFB, % (*n*/ 61 tokens)

S/N	75.4 (46)	62.3 (38)	4.9 (3)	0.0 (0)	8.2 (5)
M	14.8 (9)	13.1 (8)	0.0 (0)	0.0 (0)	1.6 (1)
W/O	9.8 (6)	--	6.6 (4)	0.0 (0)	3.3 (2)
NOG	0.0 (0)	75.4 (46)	11.5 (7)	0.0 (0)	13.1 (8)
n=	61	**≠RL**	**R**	**R+L**	**L**

JMB, % (*n*/ 46 tokens)

S/N	21.7 (10)	10.9 (5)	4.3 (2)	0.0 (0)	6.5 (3)
M	19.6 (9)	6.5 (3)	8.7 (4)	0.0 (0)	4.3 (2)
W/O	54.3 (25)	--	52.2 (24)	0.0 (0)	2.2 (1)
NOG	4.3 (2)	17.4 (8)	65.2 (30)	0.0 (0)	13.0 (6)
n=	46	**≠RL**	**R**	**R+L**	**L**

In terms of which auditory-perceptual categories were produced more often in this environment versus each speaker's overall averages (Table 28), all speakers produced higher than average numbers of not only open rhotics but also mid-range and strong/narrow rhotic as well. Speaker JMB, for instance, produced higher rates for rhotics of all three stricture categories, bringing his overall percentage for the category 'rhotic' to 65.2% (*n*= 30), nearly 25% higher than his average for all environments of 40.7%.

Speaker JMA had similar increase in the category 'rhotic,' producing 60.3% (*n*= 41) following /ɯ/ versus 37.6% across all environments. JMA differed from JMB, though, in that he also produced more mid- and weak/open laterals while JMB's increase in rhotics seems to have been at the expense of laterals. For the females, both JFA and JFB had increased percentages of strong/narrow rhotics, laterals, and non-lateral non-rhotics. With JFA in particular, she produced no mid- or weak/open laterals in this environment, making for a much lower percentage for her category 'lateral' in this environment (8.0%, *n*= 2) versus her average across all environments (29.8%). JFB, for her part, produced higher percentages of rhotics (11.5%, *n*= 7; 6.0% overall) as well as more strong/narrow non-lateral non-rhotics (62.3%, *n*= 38; 55.9% overall).

4.7.5 /_e/

A total 368 of the total 1,535 tokens examined here occurred before the mid-front vowel /e/. Table 29 summarizes the transcriptions of each speaker's realizations of /r/ before /e/ which accounted for more than 5% of his or her productions. Auditory-perceptual categories for the transcriptions are summarized further below in Table 30.

Table 29. Transcriptions for /r/ before /e/ (NOG: no audible oral gesture). Values (*n*, %) apply additively to each symbol where more than is listed.

All speakers (*n*= 368; 100%)	JFA (*n*= 72; 100%)	JFB (*n*= 91; 100%)	JMA (*n*= 126; 100%)	JMB (*n*= 79; 100%)
ɺ̞ (*n*= 60; 16.3%) ɹ (*n*= 45; 12.2%) l̆, ɾ̞, NOG (*n/3*= 35; 9.5%) ɺ, ɾ̞ (*n/2*= 29; 7.9%) ɾ (*n*= 28; 7.6%) ɾ̥ (*n*= 20; 5.4%)	ɺ̞ (*n*= 18; 25.0%) ɾ̞ (*n*= 13; 18.1%) ɺ, ɾ (*n/2*= 10; 13.9%) ɾ̞, ɹ (*n/2*= 6; 8.3%)	ɾ̞ (*n*= 28; 30.8%) ɾ (*n*= 16; 17.6%) ɾ̥ (*n*= 15; 16.5%) ɺ̞ (*n*= 12; 13.2%) l̆ (*n*= 5; 5.5%)	ɹ (*n*= 27; 21.4%) NOG (*n*= 22; 17.5%) ɺ̞ (*n*= 19; 15.1%) l̆ (*n*= 14; 11.1%) ɺ (*n*= 13; 10.3%) ˞ɺ (*n*= 8; 6.3%) ɾ̞ (*n*= 7; 5.6%)	l̆ (*n*= 16; 20.3%) ɹ, ɺ̞ (*n/2*= 11; 13.9%) NOG (*n*= 9; 11.4%) ɾ̞ (*n*= 6; 7.6%) ˞l̆, ɺ, ɾ̥ (*n/3*= 4; 5.1%)
others (*n*= 52; 14.1%)	others (*n*= 9; 12.5%)	others (*n*= 15; 16.5%)	others (*n*= 16; 12.7%)	others (*n*= 14; 17.7%)

Similar to other environments, no single transcription accounted for more than half of any speaker's realizations for /r/. Speaker JFA's most frequent realization was the lowered lateral flap [ɺ̞] (25%, *n*= 18) followed by the lowered non-lateral flap [ɾ̞] (18.1%,

n= 13). The most frequent realizations for JFB was the raised and mid-range non-lateral flaps [ɾ̝, ɾ] (30.8% and 17.6%; n= 28, 16). Although JMA's most frequent realization was the rhotic approximant [ɹ] (21.4%, n= 27); next-most frequently he produced no audible oral gesture 22 times, or 17.5%. For JMB, the rhotic approximant [ɹ] was second-most frequent along with the lowered lateral flap [ɺ̞], each of which occurred 11 times or 13.9%. JMB's most frequent realization for /r/ before /e/ was the short lateral approximant [l̆] (20.3%, n= 16).

Table 30. Category frequencies before /e/ (S/N: strong/narrow, M: mid-range, W/O: weak/open; ≠RL: non-rhotic non-lateral, R: rhotic, R+L: rhotic plus lateral, L: lateral; NOG: no audible oral gesture)

JFA, % (n/ 72 tokens)					
S/N	16.7 (12)	12.5 (9)	0.0 (0)	0.0 (0)	4.2 (3)
M	30.6 (22)	13.9 (10)	1.4 (1)	1.4 (1)	13.9 (10)
W/O	51.4 (37)	--	26.4 (19)	0.0 (0)	25.0 (18)
NOG	1.4 (1)	26.4 (19)	27.8 (20)	1.4 (1)	43.1 (31)
n=	72	≠RL	R	R+L	L

JMA, % (n/ 126 tokens)					
S/N	21.4 (27)	0.8 (1)	0.8 (1)	5.6 (7)	14.3 (18)
M	18.3 (23)	0.8 (1)	0.8 (1)	6.3 (8)	10.3 (13)
W/O	42.9 (54)	--	27.8 (35)	0.0 (0)	15.1 (19)
NOG	17.5 (22)	1.6 (2)	29.4 (37)	11.9 (15)	39.7 (50)
n=	126	≠RL	R	R+L	L

JFB, % (n/ 91 tokens)					
S/N	58.2 (53)	50.5 (46)	2.2 (2)	0.0 (0)	5.5 (5)
M	20.9 (19)	17.6 (16)	1.1 (1)	0.0 (0)	2.2 (2)
W/O	17.6 (16)	--	4.4 (4)	0.0 (0)	13.2 (12)
NOG	3.3 (3)	68.1 (62)	7.7 (7)	0.0 (0)	20.9 (19)
n=	91	≠RL	R	R+L	L

JMB, % (n/ 79 tokens)					
S/N	38.0 (30)	7.6 (6)	1.3 (1)	5.1 (4)	24.1 (19)
M	12.7 (10)	1.3 (1)	2.5 (2)	3.8 (3)	5.1 (4)
W/O	38.0 (30)	--	24.1 (19)	0.0 (0)	13.9 (11)
NOG	11.4 (9)	8.9 (7)	27.8 (22)	8.9 (7)	43.0 (34)
n=	79	≠RL	R	R+L	L

In terms of auditory-perceptual categories, the results in Table 30 suggest an overall drift away from non-rhotic non-laterals to rhotic and/or lateral categories. JFA produced rhotics and non-rhotic non-laterals of all stricture types, and along with a single instance of a combined rhotic and lateral token, the bulk of her tokens were categorized as either mid-range or weak/open laterals. In all, laterals accounted for 43.0% of her tokens before /e/ as compared to her average for this category across all environments, 29.8%. Speaker JMB patterned similarly with JFA in that his proportions of lateral and rhotic-plus-lateral categories were higher than his average values (43.0% and 8.9% before /e/ versus 34.0% and 5.0% overall). Again like JFA, the increase in JMB's rhotic + lateral and lateral categories seem to be at the expense of his rhotic and non-rhotic non-lateral categories. Speaker JMA's results are comparable to JMB's in that he also produced a higher proportion of rhotic + laterals (11.9% versus 3.8% overall). Although JMA produced more weak/open laterals (15.1%, $n = 19$, versus 9.9%), his percentage of laterals of all stricture types was slightly lower before /e/ than across all environments (39.7% versus 40.4%), due to fewer productions of mid-range and strong/narrow laterals (10.3% and 14.3% versus 11.4% and 19.0%). Also, as was the case for JFA and JMB, JMA produced fewer instances of both the rhotic and the non-rhotic non-lateral categories (1.6% and 29.4% before /e/ versus 6.7% and 37.6% overall). For her part, the bulk of JFB's tokens were characteristically mid-range and strong/narrow non-rhotic non-laterals, which accounted for 68.1% (*n*= 62) of her tokens. This figure is lower than her overall average of 74.3%, however, and the decrease in this category before /e/ was realized as increases in her proportions of rhotics of all stricture types (7.7% from 6.0%) and laterals (20.9% up from 18.4%).

4.7.6 /e_/

One-hundred seventy four tokens of /r/ occurred after the vowel /e/; transcriptions of the 174 are summarized in Table 31, while their perceptual categories are summarized further below in Table 32.

Table 31. Transcriptions for /r/ following /e/ (NOG: no audible oral gesture). Values (*n*, %) apply additively to each symbol where more than is listed.

All speakers (*n*= 174; 100%)	JFA (*n*= 34; 100%)	JFB (*n*= 48; 100%)	JMA (*n*= 51; 100%)	JMB (*n*= 41; 100%)
ɾ (*n*= 37; 21.3%)	ɾ̞ (*n*= 9; 26.5%)	ɾ̞ (*n*= 20; 41.7%)	ɺ (*n*= 11; 21.6%)	ɹ (*n*= 9; 22.0%)
ɾ̞ (*n*= 36; 20.7%)	ɾ (*n*= 8; 23.5%)	ɾ (*n*= 14; 29.2%)	ɾ, ɾ̝ (*n/2*= 8; 15.7%)	ɾ (*n*= 7; 17.1%)
ɺ (*n*= 21; 12.1%)	ɺ (*n*= 7; 20.6%)	ɾ̥, ɺ̝ (*n/2*= 4; 8.3%)	ɹ (*n*= 5; 9.8%)	ɾ̞ (*n*= 5; 12.2%)
ɾ̝ (*n*= 19; 10.9%)	ɾ̝ (*n*= 5; 14.7%)	ɾ̝ (*n*= 3; 6.2%)	l, ɺ̝, NOG (*n/3*= 4; 7.8%)	ɺ̆, ɺ, ɺ̝, ɾ̝ (*n/4*= 3; 7.3%)
ɹ (*n*= 15; 8.6%)	ɺ̝, l (*n/2*= 2; 5.9%)		ɺ̆ (*n*= 3; 5.9%)	
ɺ̝ (*n*= 13; 7.5%)				
others (*n*= 33; 19.0%)	others (*n*= 1; 2.9%)	others (*n*= 3; 6.2%)	others (*n*= 4; 7.8%)	others (*n*= 8; 19.5%)

The most frequently produced variant for both female speakers (JFA, JFB) was the raised non-lateral flap [ɾ̝] (JFA: 26.5%, *n*= 9; JFB: 41.7%, *n*= 20) followed by the mid-range non-lateral, or 'plain' non-lateral flap [ɾ] (JFA: 23.5%, *n*= 8; JFB: 29.2%, *n*= 14). JMA's most frequent realization for /r/ in this environment was the lateral flap [ɺ] (21.6%, *n*= 11) followed by 8 instances (15.7%) of each of the non-lateral flap [ɾ] and the lowered non-lateral flap [ɾ̞]. The rhotic approximant [ɹ] was JMB's most frequent realization at 22.0% (*n*= 9), followed by the mid-range non-lateral flap at 17.1% (*n*= 7).

Table 32. Category frequencies following /e/ (S/N: strong/narrow, M: mid-range, W/O: weak/open; ≠RL: non-rhotic non-lateral, R: rhotic, R+L: rhotic plus lateral, L: lateral; NOG: no audible oral gesture)

JFA, % (*n*/ 34 tokens)

	n= 34	≠RL	R	R+L	L
S/N	32.4 (11)	26.5 (9)	0.0 (0)	0.0 (0)	5.9 (2)
M	44.1 (15)	23.5 (8)	0.0 (0)	0.0 (0)	20.6 (7)
W/O	23.5 (8)	--	17.6 (6)	0.0 (0)	5.9 (2)
NOG	0.0 (0)	50.0 (17)	17.6 (6)	0.0 (0)	32.4 (11)

JMA, % (*n*/ 51 tokens)

	n= 51	≠RL	R	R+L	L
S/N	19.6 (10)	5.9 (3)	0.0 (0)	0.0 (0)	13.7 (7)
M	39.2 (20)	15.7 (8)	2.0 (1)	0.0 (0)	21.6 (11)
W/O	33.3 (17)	--	25.5 (13)	0.0 (0)	7.8 (4)
NOG	7.8 (4)	21.6 (11)	27.5 (14)	0.0 (0)	43.1 (22)

JFB, % (*n*/ 48 tokens)

	n= 48	≠RL	R	R+L	L
S/N	54.2 (26)	54.2 (26)	0.0 (0)	0.0 (0)	0.0 (0)
M	31.2 (15)	29.2 (14)	2.1 (1)	0.0 (0)	0.0 (0)
W/O	14.6 (7)	--	6.2 (3)	0.0 (0)	8.3 (4)
NOG	0.0 (0)	83.3 (40)	8.8 (4)	0.0 (0)	8.8 (4)

JMB, % (*n*/ 41 tokens)

	n= 41	≠RL	R	R+L	L
S/N	26.8 (11)	19.5 (8)	0.0 (0)	0.0 (0)	7.3 (3)
M	29.3 (12)	17.1 (7)	4.9 (2)	0.0 (0)	7.3 (3)
W/O	41.5 (17)	--	29.3 (12)	2.4 (1)	9.8 (4)
NOG	2.4 (1)	36.6 (15)	34.1 (14)	2.4 (1)	24.4 (10)

The categorical results summarized in Table 32 suggest some common patterns across the four speakers. All four speakers produced higher percentages of mid-range non-rhotic non-laterals, and all but speaker JFB similarly produced higher percentages of the strong/narrow non-rhotic non-lateral category. Mid-range stricture types were also more frequent for all speakers, although their realizations differed for the speakers in terms of the laterality-versus-rhoticity dimension. JFA had higher proportions of both mid-range and strong/narrow non-rhotic non-laterals (23.5% and 26.5% after /e/ versus 15.9% and 20.3% overall) as did speakers JMA and JMB (JMA: 15.7% and 5.9% versus 5.2% and 1.5%; JMB: 17.1% and 19.5% versus 5.1% and 8.8%). JFA and JMA also had higher proportions of mid-range laterals compared with their overall averages for the same category (JFA: 20.6% after /e/ versus 10.2% overall; JMA: 21.6% versus 11.4%). Moreover, and in spite of lower percentages for laterals of other stricture types, JFA's and JMA's higher proportions of mid-range laterals made for higher percentages for their overall 'lateral' categories (32.4% versus 29.8% for JFA, and 43.1% versus 40.4% for JMA). While JFB and JMB had higher proportions of weak/open laterals (8.3% and 9.8%, $n = 4$ before /e/ versus 7.3% and 8.5% overall), both speakers' percentages for the general 'lateral' category were lower following /e/ (JFB: 8.8% versus 18.4%; JMB: 24.4% versus 34.0%). JFB also had a slightly higher proportion of rhotics (8.8%) compared with her overall average

(6.0%), and inspite of a lower proportion of strong/narrow non-rhotic non-laterals (54.2% versus 55.9%), her percentage for the general category ‘non-rhotic non-lateral’ was higher than her overall average (83.3% versus 74.3%) due to her higher proportion of mid-range realizations.

4.7.7 /_o/

A total 101 of the total 1,535 tokens examined here occurred before the vowel /o/. Table 33 summarizes the transcriptions for each speaker’s realizations of /r/ before /o/ which accounted for more than 5% of his or her productions; further below, Table 34 provides a summary of the auditory-perceptual categories to which the transcriptions belong.

Table 33. Transcriptions for /r/ before /o/. Values (*n*, %) apply additively to each symbol where more than is listed.

All speakers (*n*= 101; 100%)	**JFA (*n*= 15; 100%)**	**JFB (*n*= 19; 100%)**	**JMA (*n*= 40; 100%)**	**JMB (*n*= 27; 100%)**
ɹ (*n*= 24; 23.8%) ɾ̝ (*n*= 17; 16.8%) ɾ̞ (*n*= 13; 12.9%) ɾ, l̆ (*n/2*= 9; 8.9%) ɺ (*n*= 8; 7.9%) ɾʴ (*n*= 5; 5.0%)	ɾ̝ (*n*= 5; 33.3%) ɾ̞ (*n*= 4; 26.7%) ɹ (*n*= 3; 20.0%) ɾ̝ʴ (*n*= 2; 13.3%) ɾ (*n*= 1; 6.7%)	ɾ̝ (*n*= 10; 52.6%) ɾ, l̆ (*n/2*= 2; 10.5%) ɾ̥, d, ˀɾ̝, ɾ, ɾ̞ (*n/5*= 1; 5.3%)	ɹ (*n*= 11; 27.5%) ɾ̞, ɺ (*n/2*= 6; 15.0%) ɾ (*n*= 5; 12.5%) ɾʴ, l̆ (*n/2*= 4; 10.0%)	ɹ (*n*= 10; 37.0%) l̆ (*n*= 3; 11.1%) ɺ, l, ɾ̞, ɾ̝ (*n/4*= 2; 7.4%)
others (*n*= 16; 15.8%)	others (*n*= 0; 0.0%)	others (*n*= 0; 0.0%)	others (*n*= 17; 13.5%)	others (*n*= 6; 22.2%)

In terms of the variants each speaker produced, both male speakers' most common realization was the rhotic approximant [ɹ] (JMA: 27.5%, *n*= 11; JMB: 37.0%, *n*= 10). JMA's next most frequent (15.0%, *n*= 6) realizations were the lowered non-lateral flap [ɾ̞] and mid-range lateral flap [ɺ]. JMB's second-most frequent realization was the short lateral approximant [l̆], which occurred 3 times, or 11.1%. The same variant was also JFB's second-most frequent variant along with the plain non-lateral flap, each of which occurred twice (10.5%). Similar to her productions in other phonological environments, JFB's most frequent realization for /r/ before /o/ was the raised non-lateral flap [ɾ̝], which she produced 10 times, or 52.6%. JFA's 15 pre-/o/ tokens were divided chiefly amongst raised non-lateral flaps (5, or 33.3%), lowered non-lateral flaps (4, or 26.7%), and rhotic approximants [ɹ] (3, or 20.0%),

As for category results (Table 34), strong claims based on proportional differences between results before /o/ and each speaker's overall averages are somewhat difficult to make considering the low numbers of tokens for the two female speakers. That said, JFA and JFB both have higher proportions of rhotics in general (JFA: 60.0% versus 30.5%, JFB: 15.8% versus 6.0%) as well as strong/narrow non-laterals (JFA: 33.3% versus 20.3%, JFB: 63.2% versus 55.9%). Also, while JFB's two instances of strong/narrow laterals make for a higher proportion (10.5% versus 4.5% overall), there seems to be a shift here away from laterality on the whole. JFA, for instance, produced no laterals here, despite laterals accounting for 29.8% of her productions in all categories.

Table 34. Category frequencies before /o/(S/N: strong/narrow, M: mid-range, W/O: weak/open; ≠RL: non-rhotic non-lateral, R: rhotic, R+L: rhotic plus lateral, L: lateral; NOG: no audible oral gesture)

JFA, % (*n*/ 15 tokens)

S/N	46.7 (7)	33.3 (5)	13.3 (2)	0.0 (0)	0.0 (0)
M	6.7 (1)	6.7 (1)	0.0 (0)	0.0 (0)	0.0 (0)
W/O	46.7 (7)	--	46.7 (7)	0.0 (0)	0.0 (0)
NOG	0.0 (0)	40.0 (6)	60.0 (9)	0.0 (0)	0.0 (0)
n =	15	**≠RL**	**R**	**R+L**	**L**

JMA, % (*n*/ 40 tokens)

S/N	12.5 (5)	0.0 (0)	0.0 (0)	0.0 (0)	12.5 (5)
M	42.5 (17)	12.5 (5)	12.5 (5)	2.5 (1)	15.0 (6)
W/O	45.0 (18)	--	42.5 (17)	0.0 (0)	2.5 (1)
NOG	0.0 (0)	12.5 (5)	55.0 (22)	2.5 (1)	30.0 (12)
n =	40	**≠RL**	**R**	**R+L**	**L**

JFB, % (*n*/ 19 tokens)

S/N	78.9 (15)	63.2 (12)	5.3 (1)	0.0 (0)	10.5 (2)
M	15.8 (3)	10.5 (2)	5.3 (1)	0.0 (0)	0.0 (0)
W/O	5.3 (1)	--	5.3 (1)	0.0 (0)	0.0 (0)
NOG	0.0 (0)	73.7 (14)	15.8 (3)	0.0 (0)	10.5 (2)
n =	19	**≠RL**	**R**	**R+L**	**L**

JMB, % (*n*/ 27 tokens)

S/N	40.7 (11)	11.1 (3)	7.4 (2)	3.7 (1)	18.5 (5)
M	11.1 (3)	3.7 (1)	0.0 (0)	0.0 (0)	7.4 (2)
W/O	44.4 (12)	--	44.4 (12)	0.0 (0)	0.0 (0)
NOG	3.7 (1)	14.8 (4)	51.9 (14)	3.7 (1)	25.9 (7)
n =	27	**≠RL**	**R**	**R+L**	**L**

While both male speakers also had higher proportions of the 'rhotic' category (JMA: 55.0% versus 37.6%; JMB 51.9% versus 40.7%), the 'drift' in their other category results differed in terms of stricture type as opposed to laterality/rhoticity. JMA showed increased percentages of mid-range strictures (42.5% versus 24.0%) across all lateral/rhotic categories while JMB also produced, to a lesser degree, higher percentages across the laterality/rhoticity dimension, but with strong/narrow stricture types (40.7% versus 39.3%)

4.7.8 /o_/

Two hundred eighty six of the 1,535 tokens in the dataset occurred following /o/; transcriptions of the 286 are summarized in Table 35, and their corresponding auditory-perceptual categories are summarized later in Table 36.

Table 35. Transcriptions for /ɾ/ following /o/. Values (*n*, %) apply additively to each symbol where more than is listed.

All speakers (*n*= 286; 100%)	**JFA (*n*= 57; 100%)**	**JFB (*n*= 58; 100%)**	**JMA (*n*= 105; 100%)**	**JMB (*n*= 66; 100%)**
ɹ (*n*= 63; 22.0%) ɾ̠ (*n*= 22; 11.5%) ɾ̟ (*n*= 29; 10.1%) ɺ̟ (*n*= 26; 9.1%) NOG (*n*= 23; 8.0%) l̆ (*n*= 22; 7.7%) ɾ (*n*= 17; 5.9%)	ɾ̟ (*n*= 15; 26.3%) ɹ (*n*= 11; 19.3%) ɾ̠ (*n*= 8; 14.0%) ɾ (*n*= 6; 10.5%) ɺ̟ (*n*= 4; 7.0%) ɾ̠ʲ (*n*= 3; 5.3%)	ɾ̠ (*n*= 21; 36.2%) ɾ̥ (*n*= 11; 19.0%) ɾ (*n*= 9; 15.5%) ɺ̟ (*n*= 5; 8.6%) d, NOG (*n/2*= 3; 5.2%)	ɹ (*n*= 38; 36.2%) NOG (*n*= 13; 12.4%) ɺ̟ (*n*= 11; 10.5%) ɾ̟ (*n*= 10; 9.5%) ɺ (*n*= 9; 8.6%) l̆ (*n*= 8; 7.6%)	ɹ (*n*= 14; 21.2%) l̆ (*n*= 12; 18.2%) ɺ, ɺ̟ (*n/2*= 6; 9.1%) NOG (*n*= 5; 7.6%) ɾ̠ (*n*= 4; 6.1%)
others (*n*= 58; 20.3%)	others (*n*= 10; 17.5%)	others (*n*= 6; 10.3%)	others (*n*= 16; 15.2%)	others (*n*= 19; 28.8%)

The somewhat disproportionate distribution of tokens among the four speakers is likely due to JMA's use of the masculine first-person pronoun *ore*, which contributed to his 105 tokens as compared to the two female speakers (JFA: 57, JFB: 58) and the other male speaker, JMB (66) who tended to use the masculine pronoun *boku* to refer to himself along with the occasional *ore*. As for how each speaker's most frequent realizations were transcribed, both male speakers' most frequent variant was the rhotic approximant [ɹ] (JMA: 36.2%, n= 38; JMB: 21.2%, n= 14). A distant second for JMA were deletions, or realizations where no oral gesture could be identified (12.4%, n= 13%). A close second for JMB was the short lateral approximant [l̆], which occurred 12 times, or 18.2%. Open rhotics were most frequent for JFA, who produced 15 (26.3%) lowered non-lateral flaps [ɾ̞] followed by 11 (19.3%) rhotic approximants [ɹ]. JFB's productions were similar to those in other environments, with raised non-lateral flaps [ɾ̝] being most frequent (36.2%, n= 21) followed by tapped fricatives [ɾ̥] and mid-range non-lateral flaps (19.0% and 15.5%, n= 11, and 9 respectively).

Table 36 displays category frequencies for /r/ variants following /o/ and compares these with each speaker's overall averages across all environments, listed earlier in Table 7. Speakers JFA and JMA's results suggest, in general, a drift away from the 'lateral' (JFA: 10.5% versus 29.8%; JMA: 25.7% versus 40.4%), and non-rhotic non-lateral (JFA: 29.8%

versus 36.2%; JMA 1.0% versus 6.7%) categories in favour of the 'rhotic' (JFA: 52.6% versus 30.5%; JMA 47.6% versus 37.6%) and 'rhotic-plus-lateral' categories.

Table 36. Category frequencies following /o/ (S/N: strong/narrow, M: mid-range, W/O: weak/open; ≠RL: non-rhotic non-lateral, R: rhotic, R+L: rhotic plus lateral, L: lateral; NOG: no audible oral gesture)

JFA, % (*n*/ 57 tokens)

S/N	24.6 (14)	19.3 (11)	0.0 (0)	1.8 (1)	3.5 (2)
M	19.3 (11)	10.5 (6)	7.0 (4)	1.8 (1)	0.0 (0)
W/O	52.6 (30)	--	45.6 (26)	0.0 (0)	7.0 (4)
NOG	3.5 (2)	29.8 (17)	52.6 (30)	3.5 (2)	10.5 (6)
n =	57	**≠RL**	**R**	**R+L**	**L**

JMA, % (*n*/ 105 tokens)

S/N	16.2 (17)	0.0 (0)	1.0 (1)	4.8 (5)	10.5 (11)
M	14.3 (15)	1.0 (1)	0.0 (0)	8.6 (9)	4.8 (5)
W/O	57.1 (60)	--	46.7 (49)	0.0 (0)	10.5 (11)
NOG	12.4 (13)	1.0 (1)	47.6 (50)	13.3 (14)	25.7 (27)
n =	105	**≠RL**	**R**	**R+L**	**L**

JFB, % (*n*/ 58 tokens)

S/N	63.8 (37)	60.3 (35)	1.7 (1)	0.0 (0)	1.7 (1)
M	17.2 (10)	15.5 (9)	1.7 (1)	0.0 (0)	0.0 (0)
W/O	13.8 (8)	--	5.2 (3)	0.0 (0)	8.6 (5)
NOG	5.2 (3)	75.9 (44)	8.6 (5)	0.0 (0)	10.3 (6)
n =	58	**≠RL**	**R**	**R+L**	**L**

JMB, % (*n*/ 66 tokens)

S/N	39.4 (26)	10.6 (7)	3.0 (2)	4.5 (3)	21.2 (14)
M	18.2 (12)	1.5 (1)	3.0 (2)	4.5 (3)	9.1 (6)
W/O	34.8 (23)	--	25.8 (17)	0.0 (0)	9.1 (6)
NOG	7.6 (5)	12.1 (8)	31.8 (21)	9.1 (6)	39.4 (26)
n =	66	**≠RL**	**R**	**R+L**	**L**

Specifically with the rhotic-plus-laterals, 2 out of JFA's total 3 productions (66.7%) of this category occurred here, following /o/. For JMA, meanwhile, 14 of his total 18 rhotic-plus-laterals (77.8%) occurred in this environment. JMB showed a similarly disproportionate number of rhotic-plus-lateral here as well: 6 of his total 15 realizations (40.0%) occurred following /o/. JMB also produced a greater proportion of laterals of all stricture types, as well as greater proportions of strong/narrow variants across the laterality/rhoticity dimension. JMB's increased percentages for the categories just mentioned came at the expense of his weak/open rhotics (25.8% versus 35.1%), mid-range rhotics (3.0% versus 3.5%), and mid-range non-rhotic non-laterals (1.5% versus 5.1%). For her part, JFB produced higher proportions of rhotics of all stricture types (8.6% versus 6.0%); of her 23 instances of realizations categorized as 'rhotic' across the entire dataset, 5 (21.7%) of them followed /o/. Although JFB produced proportionately more weak/open laterals (8.6% versus 7.3%), she had lower percentages of mid-range (0.0% versus 6.6%) and narrow (1.7% versus 4.5%) laterals. Also, while her proportion of strong/narrow non-rhotic non-lateral was higher (60.3% versus 55.9%), her proportion of mid-range non-rhotic non-laterals was slightly lower than her overall average (15.5% versus 18.4%).

4.7.9 /_a/

Four hundred sixty of the total 1,535 tokens examined here occurred before the low vowel /a/; transcriptions of variants which accounted for 5% or more of each speaker's productions are summarized in Table 37. The auditory-perceptual categories which the transcriptions represent are summarized in Table 38.

Table 37. Transcriptions for /r/ before /a/ (NOG: no audible oral gesture). Values (*n*, %) apply additively to each symbol where more than is listed.

All speakers (*n*= 460; 100%)	**JFA (*n*= 97; 100%)**	**JFB (*n*= 101; 100%)**	**JMA (*n*= 141; 100%)**	**JMB (*n*= 121; 100%)**
ɹ (*n*= 104; 22.6%)	ɾ̞ (*n*= 26; 26.8%)	ɾ̝ (*n*= 42; 41.6%)	ɹ (*n*= 41; 29.1%)	ɹ (*n*= 55; 45.5%)
ɾ̞ (*n*= 63; 13.7%)	ɾ (*n*= 14; 14.4%)	ɾ (*n*= 20; 19.8%)	ɾ̞, l̆, NOG (*n/3*= 19; 13.5%)	l̆ (*n*= 13; 10.7%)
ɾ̝ (*n*= 50; 10.9%)	ɺ (*n*= 11; 11.3%)	ɾ̥ (*n*= 10; 9.9%)	ɺ̞ (*n*= 9; 6.4%)	ɾ̞, NOG (*n/2*= 11; 9.1%)
ɾ, l̆ (*n/2*= 43; 9.3%)	ɺ̞, ɹ (*n/2*= 8; 8.2%)	ɾ̞ (*n*= 7; 6.9%)	ɺ, ɾʴ (*n/2*= 8; 5.7%)	ɺ (*n*= 8; 6.6%)
NOG (*n*= 34; 7.4%)	l̆ (*n*= 7; 7.2%)	ɺ, ɺ̞ (*n/2*= 5; 5.0%)	ɾ (*n*= 7; 5.0%)	
ɺ, ɺ̞ (*n/2*= 32; 7.0%)	l, ɾ̝ (*n/2*= 5; 5.2%)			
others (*n*= 59; 12.8%)	others (*n*= 17; 17.5%)	others (*n*= 12; 11.9%)	others (*n*= 11; 7.8%)	others (*n*= 23; 19.0%)

Similar to nearly all other pre-vocalic environments, no speaker produced any variant more than 50% of the time; however JFB's preference for raised non-lateral flaps [ɾ̝] (41.6%, *n*= 42) as well as JMB's preference for rhotic approximants (45.5%, *n*= 55) come close, with mid-range non-lateral flaps (19.8%, *n*= 20) for JFB and short lateral approximants (10.7%, *n*= 13) for JMB being somewhat distant seconds. Similar to JMB, JMA's most frequent realization for /r/ was also the rhotic approximant [ɹ], which occurred 41 times, or 29.1%. JMA's next-most frequent realizations were 19 tokens (13.5%) for each of the mid-range non-lateral flap [ɾ], short lateral approximant [l̆], and no oral gesture

(NOG). JFA produced the lowered non-lateral flap [ɾ̞] most frequently (26.8%, n= 26) followed by the mid-range non-lateral flap (14.4%, n= 14).

Category frequency results for each speaker's realizations of /r/ before /a/ are summarized in Table 38. The general tendency that is apparent in the results is, with exceptions, decreased proportions of strong/narrow strictures and overall laterality in favour of mid-range and open rhotics.

Table 38. Category frequencies before /a/ (S/N: strong/narrow, M: mid-range, W/O: weak/open; ≠RL: non-rhotic non-lateral, R: rhotic, R+L: rhotic plus lateral, L: lateral; NOG: no audible oral gesture)

JFA, % (n/ 97 tokens)

	n = 97	≠RL	R	R+L	L
S/N	23.7 (23)	10.3 (10)	1.0 (1)	0.0 (0)	12.4 (12)
M	28.9 (28)	14.4 (14)	3.1 (3)	0.0 (0)	11.3 (11)
W/O	44.3 (43)	--	35.1 (34)	0.0 (0)	9.3 (9)
NOG	3.1 (3)	24.7 (24)	39.2 (38)	0.0 (0)	33.0 (32)

JMA, % (n/ 141 tokens)

	n = 141	≠RL	R	R+L	L
S/N	19.9 (28)	0.7 (1)	0.0 (0)	0.0 (0)	19.1 (27)
M	17.0 (24)	5.0 (7)	6.4 (9)	0.0 (0)	5.7 (8)
W/O	49.6 (70)	--	43.3 (61)	0.0 (0)	6.4 (9)
NOG	13.5 (19)	5.7 (8)	49.6 (70)	0.0 (0)	31.2 (44)

JFB, % (n/ 101 tokens)

	n = 101	≠RL	R	R+L	L
S/N	60.4 (61)	53.5 (54)	2.0 (2)	0.0 (0)	5.0 (5)
M	26.7 (27)	21.8 (22)	0.0 (0)	0.0 (0)	5.0 (5)
W/O	11.9 (12)	--	6.9 (7)	0.0 (0)	5.0 (5)
NOG	1.0 (1)	75.2 (76)	8.9 (9)	0.0 (0)	14.9 (15)

JMB, % (n/ 121 tokens)

	n= 121	≠RL	R	R+L	L
S/N	14.0 (17)	1.7 (2)	0.8 (1)	0.0 (0)	11.6 (14)
M	12.4 (15)	1.7 (2)	4.1 (5)	0.0 (0)	6.6 (8)
W/O	64.5 (78)	--	55.4 (67)	0.8 (1)	8.3 (10)
NOG	9.1 (11)	3.3 (4)	60.3 (73)	0.8 (1)	26.4 (32)

JFA produced higher proportions of mid-range and weak/open rhotics here compared with her averages across all environments (35.4% and 3.1% versus 27.6% and 1.9%); half of her total 6 mid-range rhotics occurred before /a/. Unlike the other three speakers, JFA had a higher proportion of overall ‘laterality’ (33.0% before /a/ versus 29.8% over all environments), which was due to higher frequencies of mid-range and strong/narrow laterals (11.3% and 12.4% versus 10.2% and 6.7% respectively). Decreased were JFA’s proportions of both mid-range and strong/narrow non-rhotic non-laterals (14.4% and 10.3% here versus 15.9% and 20.3% overall), bringing her percentage of the general non-rhotic non-lateral category down to 24.7% before /a/ versus 36.2% across all environments. While JFB’s 15 laterals were evenly spread across stricture categories, yielding a higher than average percentage of strong/narrow laterals (5.0% versus 4.5%), her proportion of the general ‘lateral’ category was lower here compared to her average (14.9% versus 18.4%). Also slightly lower than average was JFB’s proportion of strong/narrow non-rhotic non-laterals (53.5% versus 55.9%). This was counterbalanced by slightly higher proportion of mid-range strictures of the same category (21.8% versus 18.4%). Similar to the other three speakers, JFB had higher numbers of rhotics in general (8.9% versus 6.0%); moreover, half of her dataset total 14 instances of weak/open rhotics occurred before /a/.

JMA's results show increased proportions of mid-range and weak/open rhotics (6.4% and 43.3% versus 5.0% and 32.4%), with lower than average proportions for all other categories except for strong/narrow laterals, the proportion of which was 0.1% higher than JMA's average across the entire dataset (19.1% versus 19.0%). JMA's general 'lateral' and 'non-rhotic non-lateral' categories accounted for 31.2% and 5.7% of his tokens before /a/ versus 40.4% and 6.7% for all environments. In contrast, his proportion for the category 'rhotic' was much higher here at 49.6% as compared with his overall average of 37.6%. JMB's pattern of category results are similar to JMA's in that he produced greater proportions of mid-range and weak/open rhotics (4.4% and 55.4% versus 3.5% and 35.1%) and lower than average proportions for nearly all other categories (1 of his dataset total 2 instances of the weak/open rhotic-plus-lateral category occurred before /a/). Echoing JMA, JMB's proportion for the general categories 'lateral' and 'non-rhotic non-lateral' were lower (26.4% and 3.3% versus 34.0% and 13.8%) while the general category 'rhotic' was higher here compared to his average across all environments (60.3% versus 40.7%).

4.7.10 /a_/

Six hundred ninety four of the study's 1,535 tokens of /r/ occurred following /a/; transcriptions are summarized in Table 39 while their corresponding auditory-perceptual categories are summarized further below in Table 40.

Table 39. Transcriptions for /r/ following /a/ (NOG: no audible oral gesture). Values (*n*, %) apply additively to each symbol where more than is listed.

All speakers (*n*= 694; 100%)	JFA (*n*= 156; 100%)	JFB (*n*= 175; 100%)	JMA (*n*= 183; 100%)	JMB (*n*= 180; 100%)
ɾ̝ (*n*= 97; 14.0%)	ɺ̞ (*n*= 31; 19.9%)	ɾ̝ (*n*= 72; 41.1%)	l̆ (*n*= 37; 20.2%)	ɹ (*n*= 48; 26.7%)
l̆ (*n*= 82; 11.8%)	ɾ̞ (*n*= 29; 18.6%)	ɾ (*n*= 30; 17.1%)	NOG (*n*= 28; 15.3%)	l̆ (*n*= 30; 16.7%)
ɺ (*n*= 81; 11.7%)	ɾ (*n*= 25; 16.0%)	ɺ (*n*= 23; 13.1%)	ɹ (*n*= 26; 14.2%)	ɺ (*n*= 17; 9.4%)
ɺ̞, ɹ (*n/2*= 79; 11.4%)	ɺ, ɾ̝ (*n/2*= 20; 12.8%)	ɺ̞ (*n*= 16; 9.1%)	ɺ (*n*= 21; 11.5%)	ɺ̞, ɾ̞, NOG (*n/3*= 16; 8.9%)
ɾ, ɾ̞ (*n/2*= 67; 9.7%)	l̆ (*n*= 8; 5.1%)	ɾ̥ (*n*= 9; 5.1%)	ɾ̞ (*n*= 19; 10.4%)	
NOG (*n*= 50; 7.2%)			ɺ̞ (*n*= 16; 8.7%)	
others (*n*= 92; 13.3%)	others (*n*= 23; 14.7%)	others (*n*= 25; 14.3%)	others (*n*= 36; 19.7%)	others (*n*= 37; 20.6%)

JFA's two most frequent realizations for /r/ in this environment were the lowered lateral (19.9%, *n*= 31) and non-lateral flaps (18.6%, *n*= 29). These were closely followed by their mid-range analogues with 25 (16.0%) non-lateral flaps and 20 (12.8%) mid-range lateral and raised non-lateral flaps. The raised non-lateral flap [ɾ̝] was JFB's most frequent transcription (41.1%, *n*= 72) followed by the mid-range non-lateral [ɾ] (17.1%, *n*= 30) and lateral [ɺ] (13.1%, *n*= 23) flaps. Most frequent for JMA was the short lateral approximant [l̆] (20.2%, *n*= 37); no oral gesture was detected next-most often (15.3%, *n*= 28) followed by the rhotic approximant [ɹ] (14.2%, *n*= 26). This transcription was the most frequent realization for JMB, at 26.7% (*n*= 48), followed by 30 (16.7%) short lateral approximants and 17 (9.4%) mid-range lateral flaps [ɺ].

Table 40. Category frequencies following /a/ (S/N: strong/narrow, M: mid-range, W/O: weak/open; ≠RL: non-rhotic non-lateral, R: rhotic, R+L: rhotic plus lateral, L: lateral; NOG: no audible oral gesture)

JFA, % (*n*/ 156 tokens)					
S/N	24.4 (38)	14.7 (23)	0.0 (0)	0.0 (0)	9.6 (15)
M	30.8 (48)	16.0 (25)	1.3 (2)	0.6 (1)	12.8 (20)
W/O	42.3 (66)	--	21.8 (34)	0.0 (0)	20.5 (32)
NOG	2.6 (4)	30.8 (48)	23.1 (36)	0.6 (1)	42.9 (67)
n=	156	≠RL	R	R+L	L

JMA, % (*n*/ 183 tokens)					
S/N	27.3 (50)	0.5 (1)	0.0 (0)	1.1 (2)	25.7 (47)
M	23.5 (43)	4.9 (9)	6.0 (11)	1.1 (2)	11.5 (21)
W/O	33.9 (62)	--	25.1 (46)	0.0 (0)	8.7 (16)
NOG	15.3 (28)	5.5 (10)	31.1 (57)	2.2 (4)	45.9 (84)
n=	183	≠RL	R	R+L	L

JFB, % (*n*/ 175 tokens)					
S/N	56.0 (98)	51.4 (90)	0.0 (0)	0.0 (0)	4.6 (8)
M	32.0 (56)	18.3 (32)	0.6 (1)	0.0 (0)	13.1 (23)
W/O	10.9 (19)	--	1.7 (3)	0.0 (0)	9.1 (16)
NOG	1.1 (2)	69.7 (122)	2.3 (4)	0.0 (0)	26.9 (47)
n=	175	≠RL	R	R+L	L

JMB, % (*n*/ 180 tokens)					
S/N	26.7 (48)	3.3 (6)	1.7 (3)	2.2 (4)	19.4 (35)
M	16.1 (29)	1.7 (3)	2.8 (5)	1.7 (3)	10.0 (18)
W/O	48.3 (87)	--	37.8 (68)	0.6 (1)	10.0 (18)
NOG	8.9 (16)	5.0 (9)	42.2 (76)	4.4 (8)	39.4 (71)
n=	180	≠RL	R	R+L	L

The most striking pattern noticeable in the category results in Table 40 is that, with the exception of JMA's proportion of weak/open laterals (8.7% versus 9.9% overall), all speakers had higher than average numbers of all lateral stricture categories, and correspondingly higher proportions for the general category 'lateral' (JFA: 42.9% versus 29.8%; JFB: 26.9% versus 18.4%; JMA: 45.9% versus 40.4%; JMB: 39.4% versus 34.0%). All speakers also had lower than average proportions for the general 'non-rhotic non-lateral' category: 30.8% versus 36.2% for JFA, 69.7% versus 74.3% for JFB, 5.5% versus 6.7% for JMA, and 5.0% versus 13.8% for JMB. Although JMB produced higher than average proportions of weak/open rhotics (37.8% versus 35.1%) and had a higher percentage for his general 'rhotic' category (42.2% versus 40.7%), all other speakers had lower proportions for the general 'rhotic' category (JFA: 23.1% versus 30.5%; JFB: 2.3% versus 6.0%; JMA: 31.1% versus 37.6%). JMB's results stand out from those of the other three speakers in one other respect: greater proportions for all stricture types of the rhotic-plus-lateral category. JMB produced 8 of his dataset total 15 rhotic-plus-lateral tokens in this environment.

4.8 Post-pausal/word-initial /ɾ/

Insofar as the literature on Japanese /r/ suggests that /r/ is quite different following a pause and between vowels, word-initial /r/s were not considered post-pausal if they occurred directly following the final vocalic interval of a preceding word. That is, tokens that were underlyingly word-initial but were realized as intervocalic segments in the speech stream were not coded as post-pausal /r/s. Since this was the norm in the connected speech investigated here, out of 1,535 tokens of /r/ produced by the four speakers studied here, only 17 (0.011%) of them occurred following a pause - which is a disappointing number given that one of this study's aims is to quantify what sort of /r/ was used in this position. Seventeen tokens, that is, is not a number large enough to justify any sort of definitive claims. That having been said, Table 41 summarizes the transcriptions of the 17.

Table 41. Transcriptions for post-pausal /r/. Values (*n*, %) apply additively to each symbol where more than is listed.

All speakers (*n*= 17; 100%)	JFA (*n*= 1; 100%)	JFB (*n*= 5; 100%)	JMA (*n*= 7; 100%)	JMB (*n*= 4; 100%)
l, l̆ (*n/2*= 5; 29.4%) ɾ̝ (*n*= 3; 17.6%) ɾ̝ʲ (*n*= 2; 11.6%) l̆ʱ, n (*n/2*= 1; 5.9%)	ɾ̝ʲ (*n*= 1; 100.0%)	ɾ̝ (*n*= 3; 60.0%) l̆ (*n*= 2; 40.0%)	l (*n*= 5; 71.4%) l̆, n (*n/2*= 1; 14.3%)	l̆ (*n*= 2; 50.0%) l̆ʱ, ɾ̝ʲ (*n/2*= 1; 25.0%)
others (*n*= 0; 0.0%)	others (*n*= 0; 0.0%)	others (*n*= 0; 0.0%)	others (*n*= 0; 0.0%)	others (*n*= 0; 0.0%)

With the proviso that the number of tokens precludes making any strong claims about the language in general or even these particular speakers, the results in Table 41 suggest a preference for narrow/strong strictures over any other types: all of the 17 involve narrow/strong central oral strictures. Another interesting aspect of the 17 is the absence of rhotics and the frequency of laterals, including the ‘full length’ lateral approximant [l] which numbered 5 out of speaker JMA’s total 7 tokens. Of all speakers’ tokens, laterals totalled 11, transcribed as the approximant [l], short approximant [l̆], and one short lateral approximant which occurred with breathy aspiration [l̆^{ɦ}] produced during laughter. As for non-lateral realizations, 3 of the 17 tokens were transcribed as raised non-lateral flaps [ɾ̝], 2 as raised non-lateral flaps with palatal off-glides [ɾ̝^{j}], and there was one instance of the nasal stop [n]. It is worth mentioning about the nasal that it occurred post-pausally following a dysfluency (false-start) in JMA’s speech.

4.9 Post-nasal /r/

Akamatsu (1999) and others suggest that, while lateral /r/ variants may occur in any environment, they are particularly common following /n/. It was hoped that the four speakers’ more than 60 combined minutes of extemporaneous conversation examined here would include a large number of tokens representing this environment. However, the entire dataset includes only five /r/s which directly follow /n/: two instances from each of speakers JFA and JMA, and one from speaker JMB. To make matters even less robust, the two tokens by JFA (transcribed as [ɾ̞ʰ] and [ɾ̥]) occur either during laughter or coincide with its onset. That said, the three remaining instances of /r/ from this environment are laterals, transcribed as [l̆] ($n = 1$) and [l] ($n = 2$) respectively. While this certainly hints at the possibility that the post-nasal environment is amenable to laterality, or at least strong/narrow central oral strictures, a great many more tokens are needed on which to base any firm conclusions.

4.10 Creaky voice and /r/

During the initial stages of this research, it was observed that the speakers in this dataset often ended their utterances with creaky voice quality; /r/s local to these creaky intervals, moreover, seemed to be realized as either laterals or English-like rhotic approximants, as opposed to non-lateral flaps. In order to investigate this observation further, tokens were coded so as to reflect the presence or absence of perceived creak on either side of /r/. Table 42 presents a summary of the transcriptions for /r/ where creaky voice was perceived on both sides of the consonant, 66 tokens in all. Table 43 lists results for tokens whose preceding vowels were creaky (133 tokens), and Table 44 summarizes the results where a creaky vowel followed /r/ (119 tokens).

4.10.1 Creak on both sides (V̰ _ V̰)

The data in Table 42 show that the four speakers examined here generally preferred rhotics and laterals between creaky vowels to the almost complete exclusion of non-lateral non-rhotic variants.

Table 42. Transcriptions for /r/ between creaky vowels. Values (*n*, %) apply additively to each symbol where more than is listed.

All speakers (*n*= 66; 100%)	JFA (*n*= 10; 100%)	JFB (*n*= 3; 100%)	JMA (*n*= 23; 100%)	JMB (*n*= 30; 100%)
ɹ (*n*= 16; 24.2%) l̆ (*n*= 13; 19.7%) ɺ̞, ɺ (*n/2*= 9; 13.6%) l (*n*= 6; 9.1%) ɾ̞ (*n*= 5; 7.6%)	l, l̆ (*n/2*= 3; 30.0%) ɾ̞, ɾ̞ʴ, ɺ̞, ɺ (*n/4*= 1; 10.0%)	ɺ (*n*= 2; 66.7%) ɾ̞ (*n*= 1; 33.3%)	ɹ (*n*= 7; 30.4%) l̆, ɺ (*n/2*= 4; 17.4%) ɾ̞, ɾʴ, l, ˞l, lːⁿ (*n/5*= 1; 4.3%)	ɹ (*n*= 9; 30.0%) l̆ (*n*= 6; 20.0%) ɺ̞ (*n*= 5; 16.7%) l, ɺ (*n/2*= 2; 6.7%) ˞ɺ̞, ˞l̆, ɾ̞ʴ (*n/3*= 1; 3.3%)
others (*n*= 8; 12.1%)	others (*n*= 0; 0.0%)	others (*n*= 0; 0.0%)	others (*n*= 0; 0.0%)	others (*n*= 0; 0.0%)

The rhotic approximant [ɹ] accounted for nearly a quarter (*n*= 16; 24.2%) of the 66 tokens, followed by the short lateral approximant [l̆] at roughly twenty percent (*n*= 13; 19.7%). In terms of individual speaker patterns, the two male speakers both produced /r/ in this environment more frequently than the two females; JMB, for instance, had 30 tokens versus JFB's 3. For this reason, and considering that both males preferred the rhotic approximant followed by the short lateral, the 'All speakers' column of the table is skewed toward reflecting JMA's and JMB's results. The female speakers did not produce a large number of creaky tokens; however, based on the few tokens available, they had a tendency to pattern differently than the males. JFA, for her part, seemed to prefer laterals over rhotics in that she produced 3 each of the short lateral approximant [l̆] and the longer 'English-like' lateral approximant [l]. For JFB, this environment was one of the few in which she did not produce an overwhelming majority of raised non-lateral flaps. The caveat to JFB's results here, however, is that one cannot draw any substantive conclusions from only three tokens.

4.10.2 /r/ following creak (V̰ _)

A total of 133 tokens involved a creaky vowel preceding, but not following /r/. Table 43 (overleaf) summarizes the transcriptions of /r/ variants perceived in this environment.

Table 43. Transcriptions for /r/ following a creaky vowel (NOG: no audible oral gesture). Values (*n*, %) apply additively to each symbol where more than is listed.

All speakers (*n*= 133; 100%)	JFA (*n*= 20; 100%)	JFB (*n*= 13; 100%)	JMA (*n*= 41; 100%)	JMB (*n*= 59; 100%)
ɹ, NOG (*n/2*= 22; 16.5%) l̆ (*n*= 18; 13.5%) ɺ (*n*= 11; 8.3%) ɺ̝ (*n*= 10; 7.5%) l (*n*= 9; 6.8%) ɾ̝ (*n*= 7; 5.3%)	l, NOG (*n/2*= 4; 20.0%) l̆, ɾ (*n/2*= 3; 15.0%) ɾ̝, ɾ̝ʴ (*n/2*= 2; 10.0%) ɺ̝, ɺ (*n/2*= 1; 5.0%)	ɾ̝ (*n*= 6; 46.2%) ɾ, ɺ (*n/2*= 2; 15.4%) ɺ̝, d, ɾ̝̥ʱ (*n/3*= 1; 7.7%)	ɹ, NOG (*n/2*= 9; 22.0%) l̆ (*n*= 6; 14.6%) ɺ (*n*= 4; 9.8%) ɺ̝ (*n*= 3; 7.3%)	ɹ (*n*= 13; 22.0%) l̆, NOG (*n/2*= 9; 15.3%) ɺ̝ (*n*= 5; 8.5%) ɺ (*n*= 4; 6.8%) l, ɾ̝ (*n/2*= 3; 5.1%)
others (*n*= 34; 25.6%)	others (*n*= 0; 0.0%)	others (*n*= 0; 0.0%)	others (*n*= 10; 24.4%)	others (*n*= 13; 22.0%)

Similar to the between-creaky-vowels environment, the two male speakers produced a great deal more tokens of post-creak /r/ than did the female speakers, and in particular more than JFB for whom there are only 13 tokens. Of the variants she did produce, just under half (*n* = 6; 46.2%) were raised non-lateral flaps – which is very much in line with her tendency in other phonological environments. Where they were perceived as having made an oral articulatory gesture, the remaining three speakers tended toward either mid- to narrow-laterals or open rhotics for the most part. Also noteworthy of these three speakers is their relatively frequent instances of NOG (no audible oral gesture).

4.10.3 /r/ preceding creak (_V̰)

One hundred nineteen tokens preceded (but did not follow) a creaky vowel; Table 44 summarizes the transcriptions for each speaker.

Table 44. Transcriptions for /r/ preceding a creaky vowel. Values (*n*, %) apply additively to each symbol where more than is listed.

All speakers (*n*= 119; 100%)	JFA (*n*= 19; 100%)	JFB (*n*= 13; 100%)	JMA (*n*= 45; 100%)	JMB (*n*= 42; 100%)
ɹ (*n*= 28; 23.5%) l̆ (*n*= 23; 19.3%) ɺ̞ (*n*= 13; 10.9%) ɾ̞ (*n*= 11; 9.2%) ɺ (*n*= 10; 8.4%) ɾ̝ (*n*= 9; 7.6%) l (*n*= 8; 6.7%)	l (*n*= 4; 21.1%) l̆, ɾ̞ (*n/2*= 3; 15.8%) ɾ, ɾ̝, ɺ̞ (*n/3*= 2; 10.5%) ɾ̝ʴ, ɹ, ɺ (*n/3*= 1; 5.3%)	ɾ̝ (*n*= 6; 46.2%) ɺ (*n*= 2; 15.4%) ɹ, ɺ̞, l̆, d, ɾ (*n/5*= 1; 7.7%)	ɹ (*n*= 12; 26.7%) l̆ (*n*= 11; 24.4%) ɾ̞ (*n*= 5; 11.1%) ɺ (*n*= 4; 8.9%) ɺ̞, ɾʴ (*n/2*= 3; 6.7%)	ɹ (*n*= 14; 33.3%) l̆ (*n*= 8; 19.0%) ɺ̞ (*n*= 7; 16.7%) ɺ, ɾ̞ (*n/2*= 3; 7.1%)
others (*n*= 17; 14.3%)	others (*n*= 0; 0.0%)	others (*n*= 0; 0.0%)	others (*n*= 7; 15.6%)	others (*n*= 7; 16.7%)

Similar to the other situations involving creaky voice examined here, the two male speakers' productions form the bulk of the 119. Also, except for speaker JFB, there is a general tendency to prefer open rhotics (the rhotic approximant for JMA, JMB; the lowered non-lateral flap for JFA) and mid- to narrow/strong laterals ([ɺ, l̆]) over the non-lateral non-rhotic categories of variants.

Chapter Five

DISCUSSION AND CONCLUSIONS

5.1 What speech sounds occur for /r/ in extemporaneous conversation?

At least in the extemporaneous Japanese spoken by the four individuals studied here, there is a wide array of speech sounds that occur for /r/. A telling sign of the variety of sounds is that even the most frequent sounds (Table 5, repeated below as Table 45) do not account for more than 15% of the entire dataset.

Table 45. Most frequent transcriptions for /r/

Symbol	Description	Frequency	Percent
ɾ̝	raised alveolar non-lateral flap	218	14.2
ɹ	alveolar rhotic approximant	206	13.4
ɾ̞	lowered alveolar flap	166	10.8
ɾ	alveolar non-lateral flap	161	10.5
ɺ̞	lowered alveolar lateral flap	142	9.3
ɺ	alveolar lateral flap	142	9.3
l̆	short alveolar lateral approximant	134	8.7
ɾ̥	tapped alveolar fricative	63	4.1
∅	deletion between same underlying vowels	36	2.3
l	alveolar lateral approximant	32	2.1
V^j	deletion for diphthong to [j]	31	2.0
ɾ^ɹ	alveolar flap w/rhotic off-glide	30	2.0
^{ɹ}ɺ	alveolar lateral flap w/rhotic onset	19	1.2
V^e	deletion for diphthong to [e]	15	1.0
	Others (n= <15)	140	9.1
		***T*= 1535**	**100%**

Looking just at those speech sounds that occurred more than 100 times in the dataset, we can see that the variation in /r/ encompasses lateral flaps and approximants ([ɺ, ɺ̆]), rhotic approximants ([ɹ]) and reduced (lowered) non-lateral flaps ([ɾ̞]) with an auditory quality similar to [ɹ], and flaps whose articulatory closures (and the release thereof) are akin to voiced plosives. This latter type of sound, transcribe as [ɾ̝] (raised alveolar flap) was the single most frequent realization for /r/, at 14.2% (*n* = 218) of the dataset.

Interestingly, there were fewer instances (*n* = 161, or 10.5%) of the mid-range, or 'plain' alveolar flap ([ɾ]), which is often pointed to as the liquid phoneme in Japanese (e.g. Kawakami 1977; Vance 1987; Akamatsu 1997, 2000; Amanuma et al. 2004). The articulatory difference between raised and plain alveolar non-lateral flaps is conceptualized in this thesis as a difference in the strength and quality of the closure formed by the tongue tip and the alveolar region. That is, while the tongue tip for plain alveolar non-lateral flaps touches (i.e. forms a closure) with the alveolar area of the mouth, it does so weakly and only in passing. Raised alveolar non-lateral flaps, in contrast, involve a tighter (although brief) closure that allows higher air pressure to form behind the closure which, when released, results in an audible 'burst-like' sound leading into the following vowel. Between the two possibilities of a strong, tight stricture and a weak (though complete) momentary stricture behind which little air pressure is built up, there is an intermediate possibility: taps which involve a strong/narrow central articulatory stricture behind which air pressure is likely high, if only momentarily. Here the tongue tip arguably fails to make contact with the alveolar ridge, but is rather prevented from doing so by the flow of air exiting from between the ridge and the tongue tip, a configuration that both inhibits vibration of the vocal folds at the same time as causes frication noise whose brevity is akin to the release-like burst of raised flaps. Following Jesus & Shadle (2005) who attest to a tapped fricative [ɾ̥] in European Portuguese, the symbol [ɾ̥] was also adopted here. This type of realization is similar to raised non-lateral flaps in that they involve adding a non-periodic auditory effect going into a following vowel that is very much like that of a short burst of noise associated with the release of an articulatory closure as for raised flaps; for this reason, it was categorized as a strong/narrow non-lateral non-rhotic.

5.2 Laterality and the vowels /a, e, o/

As was discussed in Chapter 2 of this thesis, the vowels /a, e, o/ have been identified as being amenable to the occurrence of lateral variants of /r/ in standard Japanese. That said, none of the literature suggests that there is any sort of absolute, or categorical demand that laterals occur adjacent to /a, e, o/; rather, what the literature suggests is that laterals are *more likely* to occur adjacent to (and especially before) these vowels. As to the four speakers investigated here, comparing each speaker's proportion of the 'lateral' auditory-perceptual category adjacent to these vowels versus their averages across all environments showed different patterns of 'drift.' That is, although a speaker may have preferred one type of non-lateral realization in a given environment, their proportion of laterality may still have been higher or lower depending on the adjacency of /a, e, o/. The following sections review the four speakers' auditory-perceptual category results and 'drift' with respect to /a, e, o/.

5.2.1 Laterality before and after /a/

There was variation in the four speakers' proportions of laterality before /a/. This variation ranged from a reduced overall proportion of laterality including fewer laterals of all stricture types (JMB), reduced overall laterality but increased proportions of strong/narrow laterals (JFB, JMA), and increased overall laterality (JFA). Speaker JFA had a higher than average proportion of laterals in general (33.0% versus 29.8%), and this overall increase reflected higher than average proportions of mid-range and strong/narrow laterals (11.3% and 12.4% versus 10.2% and 6.7%). JFA did, however, produce fewer weak/open laterals (9.3% versus 13.0%). While JFA had increased laterality before /a/, speakers JFB and JMA had decreased proportions of overall laterality (JFB: 14.9% versus

18.4%; JMA: 31.2% versus 40.4%) but, similar to JFA, increased proportions of strong/narrow laterals (JFB: 5.0% versus 4.5%; JMA: 19.1% versus 19.0%). Speaker JMB, for his part, had decreased proportions of laterals of all stricture types, bringing his proportion of overall laterality before /a/ to 26.4% versus 34.0% for the same category averaged across all environments.

Although the pre-/a/ environment suggests a range in the extent to which individual speakers produced laterals, there was more similarity among speakers' drift patterns when it came to the production of rhotic variants. Recall that, for the purposes of auditory-perceptual categorization used here, that 'rhotic' was defined as an auditory quality akin to that of the onset and/or offset of the Canadian English sound [ɹ], as in the words *star*, *ring*, and *starring*. Before /a/, all speakers had higher proportions of the category rhotic as compared to their averages for the same category across all environments (JFA: 39.2% versus 30.5%; JFB: 8.9% versus 6.0%; JMA: 49.6% versus 37.6%; JMB: 60.3% versus 40.7%).

Compared with /r/ before /a/, there was a more striking drift by all speakers toward increased laterality following /a/. JFA's proportion of laterals in general increased to 42.9% after /a/ versus her 29.8% average across all environments; JFB increased to 26.9% from 18.4%. For the male speakers, JMA's proportion of laterals was 45.9% following /a/ versus

40.4% overall, and JMB's proportion increased to 39.4% from 34.0%. JMB also had higher proportions for the general categories 'rhotic' (42.2% versus 40.7%) and 'rhotic-plus-lateral' (4.4% versus 4.0%). As to this latter category, 8 out of JMB's dataset total 15 tokens occurred following /a/.

The answer to the question of whether /a/ is amenable to laterality is thus mixed: while it does appear to be the case that the four speakers studied here produce more laterals following /a/, before /a/ there is greater individual variation. Moreover, if the pre-/a/ environment is to be characterized as being amenable to anything, that would be the production of [ɹ]-like sounds.

5.2.2 Laterality before and after /e/

Three out of this study's four speakers produced higher than average proportions of the general 'lateral' category before /e/. JFA's percentage here was 43.1% versus her average of 29.8% across all environments. For JFB, this was 20.9% before /e/ versus 18.4% overall, and for JMB this was 43.0% versus 34.0%. Only speaker JMA had a lower proportion for the general 'lateral' category (39.7% versus 40.4%), although his proportion of weak/open laterals was higher than his dataset average (15.1% versus 9.9%). Also noteworthy among the four speakers' results was that the pre-/e/ environment figured in higher proportions of the rhotic-plus-lateral category for both male speakers (11.9% versus 3.9% for JMA; 8.9% versus 4.0% for JMB). Proportions aside, 15 of JMA's dataset total 18 tokens of this category occurred before /e/ while 7 of JMB's total 15 tokens occurred here.

The results for laterality following /e/ were more mixed. Two speakers (JFA and JMA) had higher proportions of mid-range laterals (JFA: 20.6% versus 10.2%; JMA: 21.6% versus 11.4%) as well as higher proportions for the general category 'lateral' irrespective of stricture type (JFA: 32.4% versus 29.8%; JMA: 43.1% versus 40.4%). The other two speakers, JFB and JMB, both had slightly higher than average proportions of weak/open laterals (JFB: 8.3% versus 7.3%; JMB: 9.8% versus 8.5%) but lower than average proportions for the general 'lateral' category (JFB: 8.8% versus 19.4%; JMB: 24.4% versus 34.0%).

With respect to laterality and /e/, then, while three out of four of the speakers' higher proportions of laterals before /e/ suggest some amenability to laterality, insofar as individual speakers differed in their patterns of drift, we can only say conclusively that the production of laterals more often than not depends on the individual. Also, and with the proviso that two speakers' results cannot be generalized over an entire population, the pre-/e/ environment seems amenable to (males') higher than average production of variants that can be categorized as both rhotic and lateral. As to this last point, work that involves surveying far more people is needed in order to come to any conclusions about gender and the distribution of this category of variant.

5.2.3 Laterality before and after /o/

The categorical results for /r/ before /o/ suggested that, as opposed to laterals, the four speakers in this study produced higher than average proportions of rhotics. All four speakers, that is, had higher proportions of the general 'rhotic' category (JFA: 60.0% versus 30.5%; JFB: 15.8% versus 6.0%; JMA: 55.0% versus 37.6%; JMB: 51.9 versus 40.7%) and lower proportions of the general 'lateral' category (JFA: 0.0% versus 29.8%; JFB: 10.5% versus 18.4%; JMA: 30.0% versus 40.4%; JMB: 25.9 versus 34.0%). A proviso to these proportional figures is, however, that token numbers were quite low compared with other environments: JFA and JFB produced only 15 and 19 tokens of /r/ before /o/, while JMA had 40 and JMB had 27. Also, while all speakers produced fewer laterals overall, speakers JFB and JMB had higher proportions of strong/narrow laterals while JMA had a higher than average proportion of mid-range laterals.

Following /o/, a pattern roughly similar to that preceding /o/ was apparent. Three of the four speakers (JFA, JFB, and JMA) produced higher than average proportions for the general 'rhotic' category and lower proportions of the general category 'lateral.' For the category ‘rhotic,’ JFA had 52.6% following /o/ compared to her 30.5% average across all environments; JFB had 8.6% versus 6.0%, and JMA’s proportion for rhotic following /o/ was 47.6% compared to 37.6% across all environments. Again, proportions for ‘lateral’ were lower for these three speakers: 10.5% versus 29.8% for JFA, 10.3% versus 18.4% for JFB, and 25.7% versus 40.4% for JMA. In spite of their lower proportions for the general lateral category, it bears mention that JFB and JMA had slightly higher percentages of weak/open laterals (JFB: 8.6% versus 7.3%; JMA: 10.5% versus 9.9%). In contrast to the three speakers discussed above, JMB produced a lower than his average proportion for ‘rhotic’ (31.8% versus 40.7%) and a higher proportion for the ‘lateral’ category (39.4% versus 34.0%). An additional observation to make about the post-/o/ environment is that three of the four speakers (JFA, JMA, and JMB) produced a substantial amount of their dataset totals of the rhotic-plus-lateral category here (JFB produced no instances of this category in the dataset). Two of JFA’s total 3 tokens, 14 of JMA’s 18, and 6 of JMB’s dataset total 15 realizations of this category occurred following /o/.

In terms of what this means with respect to the question of whether /o/ is amenable to laterality, it seems that (like other environments) laterality is a possibility that can be exploited on an individual basis (as with JMB). That said, however, if we were to make any claim about /o/ it would be that it is more amenable to the production of rhotic variants as opposed to laterals.

5.3 Post-pausal and post-nasal /r/

The results of this study unfortunately did not say a lot about variants of /r/ that occur following pauses and nasals; this was due to surprisingly low numbers of tokens representing either environment. Excluding word-initial /r/s which were produced intervocalically as a function of connected speech, truly post-pausal /r/ occurred only 17 times out of the 1,535 tokens examined here. Even fewer were instances of /r/ following nasals: only 5.

With the proviso of low token numbers in mind, the post-pausal and post-nasal results do suggest hypotheses which warrant future investigation. Specifically, all of the 17 post-pausal tokens involved strong/narrow central oral strictures, with no instances of the weak/open rhotic category even though the rhotic approximant [ɹ] was produced frequently elsewhere by at least three of the four speakers (JFA, JMA, and JMB). Instead, lateral approximants ([l, l̆]) as well as raised non-lateral flaps with and without palatal off-glides ([ɾ̝, ɾ̝ʲ]) figured most prominently. The five post-nasal were similarly divided amongst the strong/narrow non-rhotic non-lateral, and strong/narrow lateral categories; however whether or not this was by chance is a question for future research to address.

5.4 Inter-speaker variation and similarity

Had the four speakers all patterned similarly with respect to which variants of /r/ they produced and when, there would be much fewer questions for future research to pursue. That said, there was similarity among the two male speakers in terms of their most frequently produced variant categories. Over a third of JMA's (40.4%) and JMB's (34.0%) realizations were categorized as laterals, with a descending order of preference for strong/narrow, mid-range, and finally weak/open central oral strictures (JMA: 19.0, 11.4, and 9.9%; JMB: 16.5, 9.0, and 8.5%). Both males also produced analogous proportions of the general rhotic category (JMA: 37.6%, JMB: 40.7%), with a strong preference for weak/open central strictures (JMA: 32.4%, JMB: 35.1%) followed by mid-range (JMA: 5.0%, JMB: 3.5%) and strong/narrow ones (JMA: 0.2%, JMB: 2.1%). JMA and JMB also had similar proportions of the rhotic-plus-lateral category (3.9% and 4.0% respectively); the relatively low frequency of these variants likely reflects the transitional nature of this category as a subset of the more robustly represented rhotic and lateral categories. Both males' production of variants which were neither rhotic nor lateral was similar in that they produced substantially fewer of these sounds compared with rhotics and laterals (JMA: 6.7%, JMB: 13.8%).

In contrast to JMA's and JMB's higher proportions of rhotics and laterals, JFB showed a very distinct overall preference for strong/narrow non-rhotic non-laterals (i.e. [ɾ̝, ɾ̥]). These variants accounted for 55.9% of her data, with a further 18.4% of her variants being categorized as mid-range non-rhotic non-lateral flaps, bringing her cumulative percentage of non-rhotic non-laterals to 74.3%. While JFB did produce rhotics (6.0%) and lateral (18.1%), the proportions of these were much lower than for the two males; additionally, where the male speakers' preference for central stricture types for laterals proceeded from strong/narrow to weak/open, JFB's laterals proceeded in an inverse pattern. Weak/open laterals were most prevalent (7.3%) followed by mid-range (6.6%) and strong/narrow (4.5%) central oral stricture types. JFB's rhotics (6.0%) were mostly of the weak/open variety (3.7%), with mid-range and strong/narrow stricture types accounting for 1.0% and 1.3%

Compared to the two males and JFB, JFA's results arguably resemble elements of both patterns. Similar to the males but unlike JFB, JFA had substantial proportions of laterals as well as rhotics (29.8% and 30.5%). JFA also produced a small number (3 in all, or 1.0%) of rhotic-plus-lateral tokens. JFA's realizations of /r/ resembled JFB's in that, different from the males, she had a higher proportion of non-rhotic non-laterals (36.2%) which, like JFB's, proceeded in frequency from strong/narrow strictures (20.3%) to mid-range ones (15.9%). Also, although JFA produced more laterals than JFB, these proceeded in frequency from weak/open laterals (13.0%) to mid-range (10.2%), and finally strong/narrow (6.7%) stricture types. Recall that both JMA and JMB's laterals proceeded from strong/narrow to weak/open stricture types.

Given the scope and nature of this study, it is not possible to make strong claims as to why any individual speaker's pattern of /r/ production might or might not resemble that of any other speaker. A number of possible factors can be identified, however, that warrant closer investigation in future work. These are: dialect, gender, and speech accommodation as well as speech divergence. Explaining the four speakers' variation as a function of dialect would first entail establishing the dialects used by the speakers involved, and on a larger scale, what pattern of variation was most common to speakers of that dialect. With the present study, the three Japanese language instructors who evaluated the four speakers' dialects generally agreed that both males spoke the Osaka dialect; JFA was evaluated as also speaking the Osaka dialect but not consistently. One judge, in fact, suggested that JFA was 'selectively' augmenting her speech with features characteristic of the dialect, but otherwise using *kyoutsuugo* ('common Japanese') (Gottlieb 2005), or a form of colloquial

Japanese not associated with any region in particular. There was the least agreement as to what dialect was spoken by JFB, whose pattern of /r/ variation was quite distinctive from the other three speakers investigated here. One judge felt that JFB spoke the Nagoya dialect (a dialect distinct from Kansai); another judge felt JFB was using Standard (Tokyo) Japanese, while the third judge felt that JFB was, in fact, a native speaker of the Kansai dialect who was avoiding dialectal forms in favour of 'common Japanese.' If it is the case that both JFA and JFB were avoiding (or only selectively using) dialect forms in favour of what, for them, were 'common' forms, this raises the interesting possibility that gender plays a role in the extent to which speakers of Japanese phonetically express or suppress their dialects depending on the circumstances of their speech. For instance, it may be the case that women moderate their use of dialect forms in order to display respect for their interlocutors, thereby helping them to feel comfortable in the discourse. Men, on the other hand, may use an alternative strategy to make their interlocutors feel comfortable. That is, men (or rather the two men studied here) may avoid non-dialect forms so as to mark their solidarity with one another.

Of course, one could argue that the hypotheses described above are the products of pre-conceived notions of how gender and language relate to one another in general, and in Japanese in particular. That said, and since the JST/ATR ESP-C corpus includes different conversational pairings of the four speakers at different times in the 10 conversation cycle, there is ample room for future research to investigate how any of the four speakers' patterns might change depending on their interlocutor.

5.5 Accounting for 'drift'

This thesis has presented its results in terms of 'drift' toward or from various categories of phonetic variants (see Appendix 2 for general summary of drift in all vocalic environments). While determining the causes of variation in how each speaker realizes his or her /r/s falls outside the scope of this study, it is nonetheless worthwhile to consider this question if for no other reason than to provide hypotheses which future research may investigate in more detail. As was mentioned in the discussion of the Japanese sound inventory at the beginning of this thesis, there is only one contrastive liquid in the language. This state of affairs is unlike Canadian English (CE), for instance, which has two phonemic liquids: /ɹ, l/. In other languages such as Tamil, there are five (e.g. Narayanan et al. 1999). It can be argued that the consequence of having one versus several liquids is that there is greater or less phonological 'space' in which variation can take place. That is, in a language like CE with more than one liquid, a speaker may be less inclined to produce a lateral variant for /ɹ/ because this could potentially be confused with /l/ by his or her interlocutor:

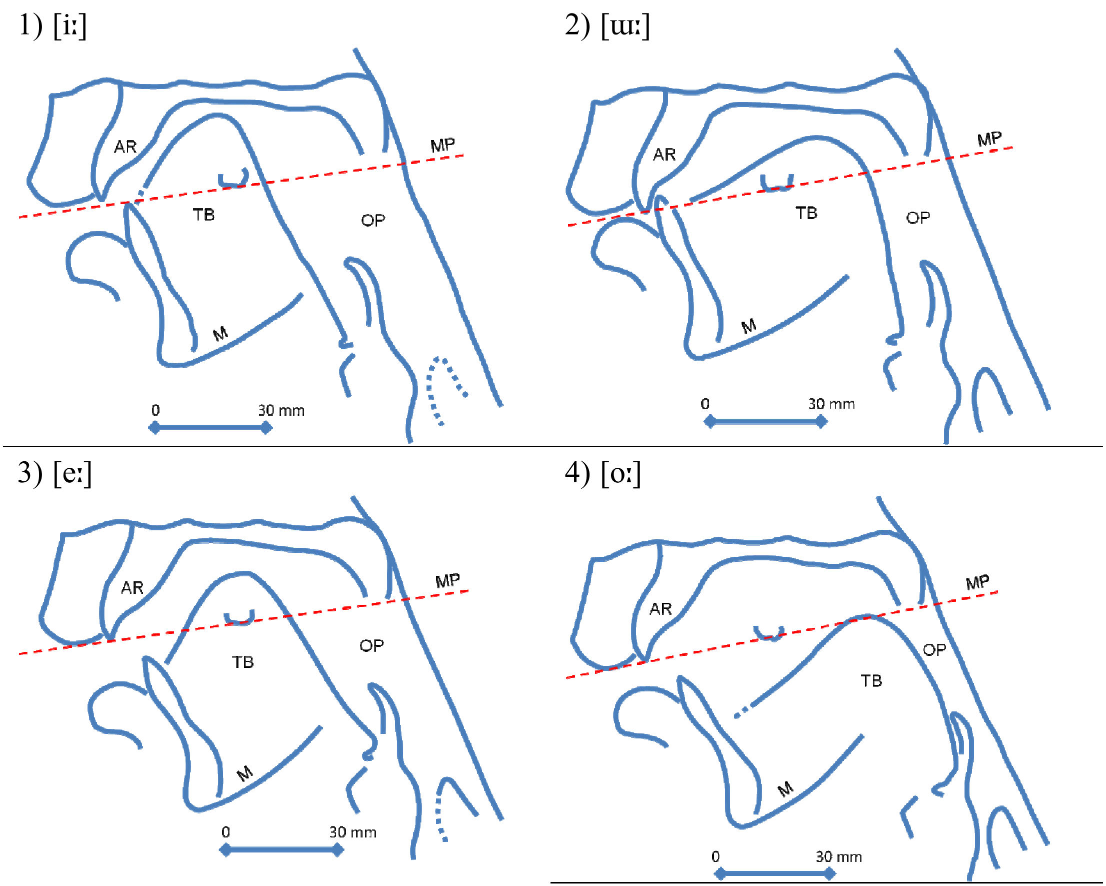
the contrast between /ɹ/ and /l/ would be neutralized (e.g. Odden 2005:68), giving rise to confusion between words like *right* and *light*. In this way, we can say that a CE speaker has less phonological space in which to vary in his or her productions of liquids. By extension, a Tamil speaker's liquid variation is even more tightly constrained. In Japanese, however, a single liquid means that speakers may vary their realizations freely without the risk of their interlocutors' confusing their sounds for any others which might otherwise distort the speaker's intended meaning.

Having established this notion of a phonological 'space' in which variation can occur, the next question is, given a wide allowable range, why do we see realizations 'drifting' towards one or another category type in certain vocalic environments. Following /o/ and preceding /a/, for instance, the four speakers studied here produced higher than average proportions of [ɹ]-like sounds. Following /a/, the four speaker's variations drifted towards increased laterality while different patterns of laterality were seen adjacent to /e/. Assuming that patterns of drift adjacent to certain vowels are not the products of conscious effort on the speakers' part, and not constrained by the danger of neutralizing any phonological contrasts with other liquids, one is left to consider the possibility that drift is the result of physiological, or kinematic factors[2].

Specifically, this idea holds that the configuration of the tongue and vocal tract for different vowels adjacent to /r/ create predispositions to certain phonetic realizations over, but not to the exclusion of, others. This proviso about non-exclusivity is important in that, if it were the case that the physiological configurations for vowels *determined* the kinds of realizations that speakers produced, there would be no inter-speaker variability. In other words, drift as a product of vowel physiology ought best to be thought of as an easily-flouted tendency and not an overriding biological imperative.

What, then, can be explained through vocal tract configurations for and between the five vowels of Japanese? As can be seen in Kokken's (1990) X-ray tracings of the vocal tract for the vowels /i, ɯ, e, o, a/ (Figure 15), the height and retractedness of the tongue body varies among the five vowels. If we assume that the production of Japanese /r/, in turn, involves a rapid articulatory gesture such that the tongue tip ideally makes a brief contact with the alveolar ridge, this creates the potential for situations where the successful completion of the gesture for /r/ is either assisted or hindered by the positioning of the tongue body.

[2] Prosodic factors may have a bearing on speakers' patterns of drift. Due to limitations of space, however, the interaction of prosody and /r/ variation will be held as an issue for future acoustical research.



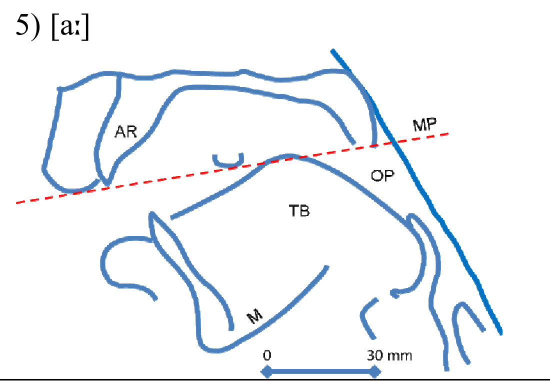


Figure 15. X-ray tracings of Japanese vowels (Kokken 1990, adapted). All images: sagittal view, facing left.

AR: Alveolar ridge
MP: Maxillary plane
TB: Tongue body
OP: Oropharynx
M: Mandible

Although this idea is similar to that of Derrick & Gick (2008) who suggested articulatory conflict determined sub-phonemic variation among North American English flaps, the key difference is that the present claim does not attempt to argue that the sub-phonemic variation in Japanese flaps is categorical. Rather, the variation is only *assisted* or *hindered* by the physiological demands placed on tongue shape called for by adjacent vowel configurations. ‘Drift,’ in turn, can be characterized as the tendency for the ballistic apical gesture for /r/ to accommodate the demands of adjacent vowels along one or a combination of three readily exploitable articulatory parameters aside from vocal fold vibration: a) stricture formation along the centre of the oral vocal tract between the blade of the tongue and the alveolar ridge, b) availability of a channel for airflow between one or both sides of the tongue body and upper molars, and/or c) modification of the shape of the pharynx by the tongue body consistent with an auditory signal akin to [ɹ]. What is meant by ‘readily exploitable’ in terms of the three above parameters is that a speaker of Japanese can increase or decrease the saliency of an /r/ variant by exploiting one or more of the parameters to varying degrees.

With respect to ‘drift’ and these parameters, the claim here is that the physiological configuration of adjacent vowels can predispose, but not determine, the exploitation of one parameter or set of parameters over another. The sections that follow develop this argument by further illustrating these parameters while making reference to the production data from the previous Chapter.

5.5.1 Alveolar stricture formation

This parameter indicates the extent to which the blade or apex of the tongue comes into contact with the alveolar region. The salience of speech sounds involving alveolar stricture formation can be modified by reducing or increasing the extent of the stricture in terms of both time and muscular strength. Strictures made by the tongue pushing tightly against the alveolar region to form complete closures, for instance, can allow for supra-glottal air pressure to increase such that vocal fold vibration is impeded and the release of the closure is accompanied by aspiration, as for [t^h]. Weaker but nonetheless complete closures are also possible, such that vocal fold vibration is not impeded and no aspiration is heard with the release of the closure, as for [d]. Strictures/closures can also be held, lengthening them temporally. If full closures involving longer durations and more strength represent one end of a spectrum along which alveolar stricture formation can be used to produce a salient speech sound, the other end of the spectrum is arguably where the vocal tract remains entirely unimpeded below the alveolar region. The tongue blade in this case is at rest behind the lower front teeth, and makes no movement towards the alveolus.

The tokens of /r/ examined in this thesis represent multiple points along the spectrum described above, ranging from a small number of voiced plosives ([d]), more frequent raised non-lateral flaps ([ɾ̝]), tapped alveolar fricatives ([ɾ̥]), non-lateral flaps ([ɾ]), and lowered (weak/open) non-lateral flaps ([ɾ̞]). In order to make a claim that the physiological demands of adjacent vowels affect the extent to which this parameter is more or less readily available to speakers, we would expect to see that, in general, vowel configurations such as that for [i] which locate the tongue blade closer to the alveolus would be more amenable to the formation of alveolar strictures. Conversely, we would expect vowels whose physiological configurations place the tongue blade further away from the alveolar region, such as for [a], to be less amenable to the formation of strong/narrow strictures. Looking again at the Kokken (1990) X-ray tracings for these vowels (repeated below as Figure 16) and patterns of drift exhibited by this study's four speakers in the environments /a_a/ and /a_i/ (Table 46), we see evidence for these hypotheses.

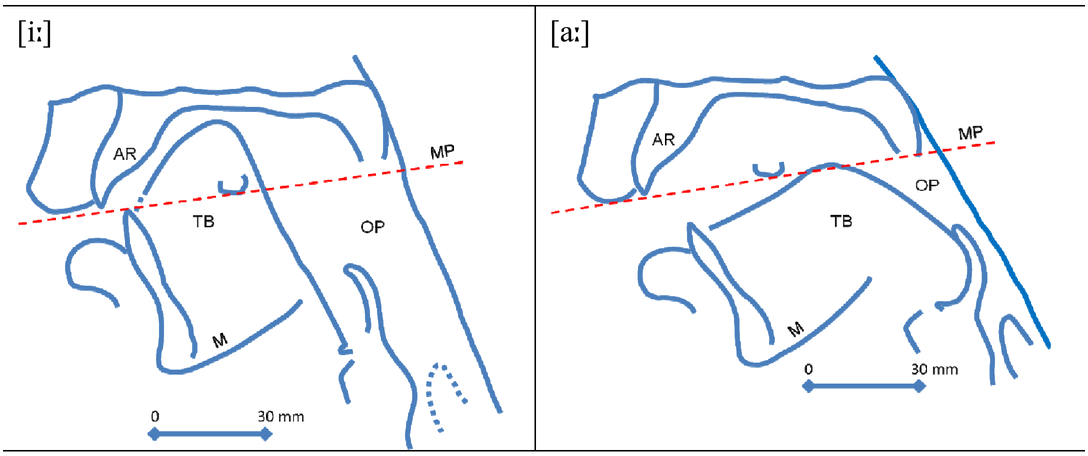


Figure 16. X-ray tracings (sagittal view, facing left) of [iː] and [aː] (Kokken 1990, adapted; AR: alveolar ridge, MP: maxillary plane, TB: tongue body, OP: oropharynx, M: mandible)

What Figure 16 shows is that the position of the tongue body (TB) for [iː] is much closer to the alveolar ridge (AR) than it is for [aː], which involves more retraction of the tongue root and lowering of the jaw, or mandible (M). The first four columns in Table 46 display each speaker's pattern of drift in the environments /a_i/ and /a_a/ relative to their dataset averages (Table 7) and with respect to the stricture-type categories defined in Chapter 3.

Table 46. Summary of category drift in /a_i/ and /a_a/ (●: higher, ○: lower, --: no change; S/N: strong/narrow, M: mid-range, W/O: weak/open, NOG: no audible oral gesture, ≠RL: non-rhotic non-lateral, R: rhotic, R+L: rhotic-plus-lateral, L: lateral)

/a_i/	S/N	M	W/O	NOG	≠RL	R	R+L	L
JFA	●	●	○	○	●	○	○	●
JFB	○	●	●	●	●	○	--	●
JMA	●	○	○	○	●	○	○	●
JMB	●	●	○	●	--	○	●	●

/a_a/	S/N	M	W/O	NOG	≠RL	R	R+L	L
JFA	○	●	●	●	○	○	○	●
JFB	○	●	●	●	○	○	--	●
JMA	●	○	○	●	●	○	○	●
JMB	○	○	●	●	○	●	○	○

Dark circles in the Table indicate a higher than average proportion for a given category while light circles indicate a lower than average proportion; two dashes (--) indicate no change. In the environment /a_i/ (see also Table 14), all speakers except for JFB produced higher than average proportions of strong/narrow (S/N) stricture types along with lower than average proportions of weak/open (W/O) stricture types. JFB patterned inversely, with a lower proportion of the S/N category and a higher proportion for W/O; considering that JFB's overall realizations of /r/ tended towards strong/narrow non-rhotic non-laterals, her pattern here is somewhat puzzling.

The four speakers' patterns of drift in the environment /a_a/ (see also Table 10) are nearly opposite to those for /a_i/. In this case, all speakers except for JMA produced lower proportions of the S/N stricture category and higher proportions of the W/O category. This environment also seemed more amenable to producing no audible oral gesture (NOG). In terms of motivating an alveolar stricture formation parameter, these data suggest that tongue position for vowels has some, if limited, influence on speakers' use of alveolar stricture formation to produce a salient speech sound. The same data will be revisited in the following sections which discuss the remaining parameters of lateral aperture and pharyngeal shape modification.

5.5.2 Lateral aperture

The availability of a channel for air to flow out of the oral cavity from around one or both sides of the tongue depends on the extent to which the tongue body forms a seal with upper teeth, or in other words, along the maxillary plane (MP). The formation of such a seal can, in turn, be manipulated by adjusting the height and/or shape of the tongue body such that air is allowed to pass through the oral cavity at the same time as a central oral stricture is maintained.

Returning to the data from the previous Section (Figure 16 and Table 46), it can be seen that, for the vowel [aː], the tongue body is below the maxillary plane, a configuration that would allow for airflow between the back of the tongue and the upper molars. Revisiting the drift pattern results in Table 46 and this time looking at the right-most column, which indicates comparative proportions for the category 'lateral,' we can see that all speakers' proportions for this category were higher in the environment /a_i/. The same is true for all but one of the speakers (JMB) in the environment /a_a/. The exception in this case, JMB's pattern, seems to have produced [ɹ]-like variants in this environment instead of laterals. Lastly with respect to laterality and the patterns of drift in these environments, it appears that stricture types need not necessarily be of the S/N variety insofar as there were many instances of mid-range as well as weak/open laterals (i.e. [ɺ, ɺ̞]) in the dataset. That said, maximally exploiting lateral aperture to produce highly salient lateral variants such as the short lateral approximant [l̆] likely involves a high degree of exploitation of alveolar stricture formation as well. That is, the central alveolar stricture provides a brace against which the muscles of the tongue can form convex shapes consistent with lateral airflow (e.g. Narayanan et al. 1997).

5.5.3 Pharyngeal modification

The idea of pharynx shape modification as an articulatory parameter to describe phonologically rhotic speech sounds stems from this author's previous work (Magnuson 2007b) which sought to interpret the primarily acoustic parameter associations for rhotics proposed by Lindau (1985) into more articulatory terms. The same work also sought to provide a model for the integration of the LPVT (laryngeal/pharyngeal vocal tract) (Esling 2005) and its conceptualization as a series of 6 valves (Edmondson & Esling 2006) into the production of phonologically rhotic speech sounds. The term *pharyngeal modification* was defined as the dynamic expansion or contraction of the pharynx via the working of some combination of three of the six valves of the LPVT (Edmondson & Esling 2006:159): laryngeal constriction (Valve 3), epiglotto-pharyngeal constriction (Valve 4), and/or pharyngeal narrowing (Valve 6) (Magnuson 2007b:1194). Insofar as the present study has made an operational distinction between rhotics as a phonological class of speech sounds and the term *rhotic* as an auditory-perceptual label for sounds that are 'akin to the sound [ɹ]' or [ɹ]-like, an addendum to the above definition here is that the shape of the pharynx be consistent with the auditory signal associated with [ɹ].

To the extent that pharyngeal modification is in this way being associated with an [ɹ]-like auditory quality, it is also important to note that pharyngeal modification is a secondary articulation that is dependent upon a concomitant primary configuration of the tongue. That is to say, without an oral lingual gesture of some kind occurring along with pharyngeal modification, the resultant speech sound would be a pharyngeal and not necessarily akin to a [ɹ]-like rhotic, *per se*. The question that remains, then, is what combination of oral lingual gesture and pharyngeal modification produces salient [ɹ]-like sounds?

Approaching this question from the [ɹ]-as-perceptual-reference point of view adopted earlier in this thesis, the tip-up /ɹ/ of North American English described by (among others) Delattre & Freeman (1964), Catford (1977), and Espy-Wilson et al. (2000) seems best suited as a candidate configuration for the [ɹ]-like sounds heard in this dataset. Specifically, and analogizing from Derrick & Gick's (2008) ultrasound study of flaps and vocalic 'r' in NAE, I propose that [ɹ]-like auditory quality in the dataset examined here is due to the co-occurrence of a linguo-alveolar flapping gesture in the oral vocal tract and a simultaneous modification of the shape of the pharynx such that an auditory quality akin to [ɹ] is produced. A prediction that follows from this hypothesis is that vowel configurations that do not involve constriction (modification) of the pharynx (e.g. [i, e]) will be less amenable to the occurrence of /r/ variants categorized auditory-perceptually categorized as rhotic.

Conversely, vowel configurations that do involve modification of the pharynx (e.g. [o, a]) will be more amenable to such variants. To examine this prediction, Figure 17 shows the Kokken (1990) X-ray tracings for the vowels [oː] and [eː] while the four speakers' patterns of drift in the environments /o_e/ and /e_/ are summarized in Table 47.

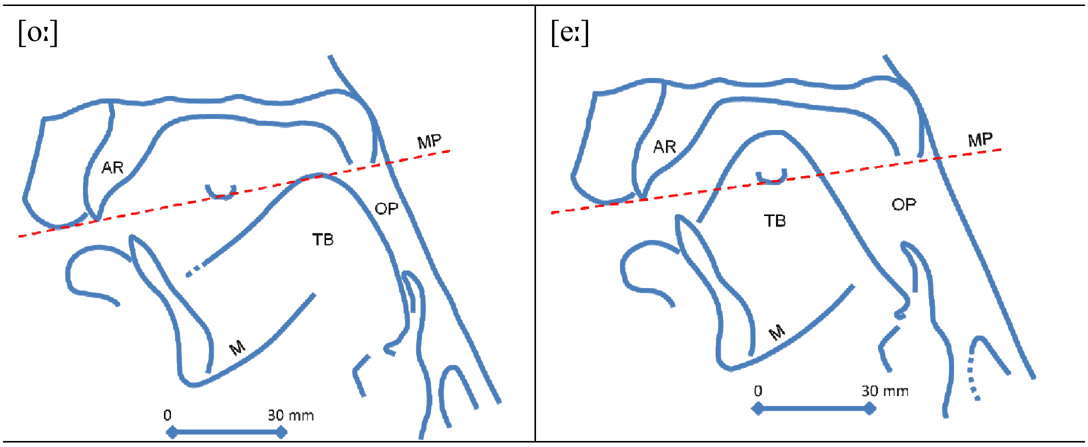


Figure 17. X-ray tracings of [oː] and [eː] (Kokken 1990, adapted; AR: alveolar ridge, MP: maxillary plane, TB: tongue body, OP: oropharynx, M: mandible)

What the X-ray tracings for [oː] and [eː] show is markedly different configurations of the tongue body (TB) both in terms of height and the extent to which the shape of the (oro-)pharynx (OP) is modified. For [eː], the OP is not constricted whereas for [oː] the OP is more narrow (Valve 6), and the root of the tongue is shown pushing the epiglottis back toward the pharyngeal wall (Valve 4). A movement by the blade of the tongue toward the alveolar ridge (AR) would involve the muscles along the top of the tongue contracting, subsequently tugging the bulge at the top of the TB forward and the tongue root and epiglottis momentarily further into the OP. The result of this collusion of gestures in turn creates the balance of acoustic trading relations (Nieto-Castanon et al. 2005) appropriate for [ɹ]. While a momentary gesture by the tongue blade or tip that does not make contact with the AR (such as a weak/open flap [ɾ̞]) would only briefly create the right circumstances for [ɹ], the salience of the [ɹ]-sound could be altered increase by holding the tongue tip/blade up via bracing between the sides of the tongue dorsum and the upper molars (or, in other words, the inverse of a lateral aperture).

Recall that the prediction made earlier was that vowel configurations such as that for [oː] would be more amenable to [ɹ]-like variants because they provide the opportunity for a tongue blade gesture to coincide with pharyngeal modification. Table 47 shows the drift patterns of the four speakers in the /o_e/ and post-/e/ environments.

Table 47. Summary of category drift in /o_e/ and /e_/ (●: higher, ○: lower, --: no change; S/N: strong/narrow, M: mid-range, W/O: weak/open, NOG: no audible oral gesture, ≠RL: non-rhotic non-lateral, R: rhotic, R+L: rhotic-plus-lateral, L: lateral)

/o_e/	S/N	M	W/O	NOG	≠RL	R	R+L	L
JFA	○	○	●	●	○	●	●	○
JFB	○	○	●	●	○	●	--	○
JMA	○	○	●	●	○	●	●	○
JMB	●	○	○	●	○	○	●	●

/e_/	S/N	M	W/O	NOG	≠RL	R	R+L	L
JFA	●	●	○	○	●	○	○	●
JFB	○	●	●	○	●	●	--	○
JMA	○	●	○	○	●	○	○	●
JMB	○	●	○	○	●	○	○	○

Again, dark circles indicate higher-than-average proportions for a given category while light circles indicate lower proportions. Looking at the column 'R' in the table, we can see that all but one speaker (JMB) produced higher than average proportions for the category 'rhotic.' Additionally, all speakers who produced instances of the rhotic-plus-lateral (R + L) category realized higher proportions of that category following /o/ (JFB produced no instances of this category in the dataset). Also noteworthy of drift in the /o_e/ environment is that the speakers who had higher proportions of rhotics also had higher proportions of the weak/open (W/O) stricture category as well as the NOG (no audible oral gesture) category. Finally, all speakers produced lower-than-average proportions of the non-rhotic non-lateral (≠RL) category in the /o_e/ environment while the opposite is true following /e/: all had higher proportions. Moreover, all but one speaker (JFB) produced lower-than-average proportions of the R category following /e/. These observations suggest that the earlier prediction about pharyngeal modification seems to be, with exceptions, correct.

5.6 Implications for the study of rhotics in general

Insofar as the phonologically rhotic (Lindau 1985, Walsh Dickey 1995) speech sounds observed in this study varied across multiple articulatory parameters, we can say that the observations generally support the idea of family resemblances. That is, rhotics form a class of sounds based on co-occurring parameters of association as opposed to any single phonetic property. In the previous sections, three parameters were proposed for relating the Japanese /r/ variants in this study: alveolar stricture formation, lateral aperture, and pharyngeal modification. This terminology represents an articulatory interpretation of an auditory signal; future research, moreover, is needed to verify through instrumental observations the extent to which this articulatory casting of the three parameters is appropriate. In a similar vein, these parameters might also be cast in acoustical terms: duration and nature of signal interruption (alveolar stricture formation), as well as the behaviour of the second (lateral aperture) and third (pharyngeal modification) formant frequencies. Insofar as the Lindau (1985) model of parameter associations is based on acoustical observations, detailed acoustical analyses of the data examined auditorily here is also needed to evaluate the extent to which the variants from this study can be accommodated by Lindau's model. In a previous re-interpretation of the family resemblances model (Magnuson 2007b), this author suggested an articulatory parameter, pharyngeal modification, as an additional means by which to co-associate rhotics in a variety of languages. This parameter, along with vocal fold vibration and aryepiglottic trilling, was motivated by the need to incorporate the functioning of the laryngeal-pharyngeal vocal tract (Esling 2005) into the primarily oral model proposed in Lindau (1985), effectively expanding its scope.

With respect to either Lindau (1985) or Magnuson (2007b), the variation presented by the four speakers in this study suggests that there is a need to further expand the scope of family resemblances to include the possibility of laterality. That is, instead of modelling the associations which bind only rhotics together (by either acoustical or articulatory parameters), it is perhaps more worthwhile to model the associations amongst liquids in general. The difference between liquids and other varieties of speech sounds, moreover, could be discussed in terms of the multiple co-occurrence of articulatory or acoustic parameters being a defining characteristic of that sound class, subject in turn to variation in degrees of saliency depending on context and individual speaker.

5.7 Summary and concluding remarks

The primary goal of this study has been to quantify variation in a phoneme known for its high degree of variation, Japanese /r/. Fifteen hundred thirty five instances of /r/ were excised from two recordings of near-natural speech data recorded by two male and two female speakers of Kansai Japanese. The realizations for /r/ spanned a wide range of variation, including raised non-lateral flaps ([ɾ̝]), tapped fricatives ([ɾ̞̊]), lowered non-lateral flaps ([ɾ̞]), short lateral as well as rhotic approximants ([l̆, ɹ]), and finally lateral and lowered lateral flaps ([ɺ, ɺ̞]). Also observed, though less frequently, were voiced plosive ([d]) realizations for /r/ as well as /r/s, such as pre-rhoticized laterals ([˞l̆, ˞ɺ]) which featured an [ɹ]-like auditory quality leading into the consonant and an [l]-like quality leading into the following vowel. Three of the speakers' /r/s could be divided among three general auditory-perceptual categories: laterals, [ɹ]-like open rhotics, and non-lateral non-rhotics. The fourth speaker (JFB) also produced some lateral and perceptually [ɹ]-like variants, but the majority of her /r/s were non-lateral flaps that also involved burst-like releases of momentary articulatory closures or frication noise produced as the result of incomplete, but narrow, closures.

When each speaker's average category frequencies were compared with their totals in specific phonological environments, patterns emerged such that the proportion of tokens of a particular category was greater or less than speakers' dataset averages. While this 'drift' towards different categories did not overshadow a given speaker's general preference for certain category types, patterns of drift did point to the influence of vowel physiology on the direction of drift in a given environment. Specifically, patterns of drift tended towards increased proportions of lateral following the retracted vowel /a/; variants categorized as auditory-perceptually rhotic ([ɹ]-like) increased proportionately following /o/.

To attempt to account for drift as well as variation within the dataset in general, three articulatory parameters were proposed: alveolar stricture formation, lateral aperture, and pharyngeal modification. It was argued that, similar to the parameters of association among rhotic speech sounds proposed by Lindau (1985), the three parameters proposed in this thesis can co-occur to bring about the myriad of different /r/ variants present in the dataset.

It is the author's hope that the enumeration of /r/ variants presented in this thesis will be of assistance in the formulation of speech production models which can robustly account for human speech as it is spoken by humans in their daily lives – complete with reductions, deletions, exaggerations, and dysfluencies. With that in mind, though, the work presented here leaves more questions than it answers. Future articulatory, acoustic, and pragmatic analysis of this dataset and ones similar to it are needed in order to gain a more robust understanding of how liquids and other speech sounds vary in natural language.

BIBLIOGRAPHY

Akamatsu, T. (1997). *Japanese Phonetics: Theory and Practice*. München, Newcastle: Lincom Europa.

Akamatsu, T. (2000). *Japanese Phonology: A Functional Approach.* München, Newcastle: Lincom Europa.

Amanuma, Y., Ohtsubo, K., Mizutani, O. (2004). *Nihongo Onseigaku* [Japanese Phonetics]. 3rd Edition. Tokyo: Kuroshiro Shuppan.

Beasley, W.G. (1972). *The Meiji Restoration*. Stanford, CA: Stanford University Press.

Boersma, P., Weenink, D. 2006. Praat: Doing phonetics by computer (Version 4.4.32) [Computer program]. Retrieved 2 Oct. 2006, from http://www.praat.org/ .

Campbell, N. (2007). Resources for Interacting with Large-Corpus Speech Data. *Oriental COCOSDA*, 4-7 Dec., Hanoi, Vietnam.

Campbell, N. (2004). Speech & expression; the value of a longitudinal corpus. *Proc. 4th International Conference on Language Resources and Evaluation*, 26-28 May, Lisbon, Portugal.

Campbell, N., (2002). Recording techniques for capturing natural everyday speech. *Proc. 3rd Language Resources and Evaluation Conference*, 29-31 May, Las Palmas, Spain.

Catford, J.C. (1977). *Fundamental Problems in Phonetics.* Bloomington, IN: Indiana University Press.

Cho, S.B. (1967). *A Phonological Study of Korean*. Uppsala: Almqvist and Wiksells. (Cited in Ladefoged & Maddieson 1996).

Delattre, P., Freeman, D.C. (1968). A dialect study of American r's by x-ray motion picture. *Linguistics* 44, 29-68.

Derrick, D., Gick, B. (2008). Quantitative analysis of subphonemic flap/tap variation in NAE. *Proc. Acoustics Week in Canada*, 5-8 Oct. 2008, Vancouver, Canada.

Edmondson, J.A., Esling, J.H. (2006). The valves of the throat and their functioning in tone, vocal register and stress: Laryngoscopic case studies. Phonology 23, 157-191.

Esling, J.H. (2005). There are no back vowels: The laryngeal articulator model. Canadian *Journal of Linguistics* 50, 13–44.

Espy-Wilson, C.Y., Boyce, S.E., Jackson, M., Narayanan, S., Alwan, A. (2000). Acoustic modeling of American English /r/. *Journal of the Acoustical Society of America* 108(1), 343-356.

Fukaya, T., Byrd, D. (2005). An articulatory examination of word-final flapping at phrase edges and interiors. *Journal of the International Phonetic Association* 35(1), 45-58.

Gottlieb, N. (2005). *Language and Society in Japan.* Cambridge: Cambridge University Press.

Grenon, I. (2005). The status of the sound [z] in Japanese. *Langues et Linguistique* 31, 63-90.

Halle, M. (1961). On the role of simplicity in linguistic descriptions. *Structure in Language and its Mathematical Aspects: Proceedings of Symposia in Applied Mathematics* 12, 89-94.

Harris, J.G. (2006). Some common r-sounds. In: J.E. Harris (Ed.). *Readings in Articulatory Phonetics: Volume 1. Consonants & Phonation Types*. Bangkok: Ek Phim Thai Co. Ltd.

Hattori, S. (1960). *Gengogaku no Houhou* [Methods in Linguistics]. Tokyo: Iwanami Shoten.

Hattori, S. (1951). *Onseigaku* [Phonetics]. Tokyo: Iwanami Shoten.

International Phonetic Association. (1999). *Handbook of the International Phonetic Association: A Guide to the Use of the International Phonetic Alphabet*. Cambridge: Cambridge University Press.

International Phonetic Association. (2005). The international phonetic alphabet (revised to 2005). [Online resource, 10 Nov. 2008 access]. Department of Linguistics, University of Victoria, http://web.uvic.ca/ling/resources/ipa/charts/IPA_chart_(C)2005.pdf.

Ito, Junko & Mester, A. (1999). The Phonological Lexicon. In John McCarthy (ed.), *Optimality Theory in phonology: A reader*. Cambridge: Blackwell. [From: Natsuko Tsujimura (ed.), *The Handbook of Japanese Linguistics*. Oxford: Blackwell, 62-100.]

JNTO [Japan National Tourist Organization]. (2007). Traditional annual events map. [Online resource, 26 Oct. 2008 access]. http://www.jnto.go.jp/eng/indepth/history/traditionalevents/index.html.

Jesus, L.M.T., Shadle, C.H. (2005). Acoustic analysis of European Portuguese uvular [χ, ʁ] and voiceless tapped alveolar [ɾ̥] fricatives. *Journal of the International Phonetic Association* 35(1), 27-44.

Joo, H. (2005). *Nihon Onseigaku Kenkyuu: Jikken Onseigaku Houhou Ronkou* [A Study on Japanese phonetics: A treatise on the research methodology of experimental phonetics]. Tokyo: Bensey Publishing.

Kawauchi, A. (1993). *Kansai-ben Tanken* [Discovering Kansai Japanese]. Osaka: Toho.

Kokken [National Language Research Institute]. (1990). *Kokuritsu Kokugo Kenkyuujo Houkoku 100. Nihongo no Boin, Shiin, Onsetsu: Chouon Undou no Jikken Onseigakuteki Kenkyuu* [Report of the National Language Research Institute no. 100.

Japanese vowels, consonants, syllables: Experimental phonetics research of articulatory movements]. Tokyo: Shuei Shuppan.

Kubozono, Haruo. (1999). Mora and syllable. In: Natsuko Tsujimura (ed.), *The Handbook of Japanese Linguistics*, pp. 31-61. Oxford: Blackwell.

Labov, W. (1966). *Social Stratification of English in New York City.* Washington, DC: Center for Applied Linguistics.

Ladefoged, P. (1964). *A Phonetic Study of West African Languages*. J.H. Greenberg & J. Spencer (eds.). West African Language Monograph Series 1. London: Cambridge University Press.

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

Levitt, H., Rabiner, L.R. (1971). Analysis of fundamental frequency contours in speech. *Journal of the Acoustical Society of America* 49(2B), 569-582.

Lindau, M. (1985). The story of /r/. In: V.A. Fromkin (Ed.), *Phonetic Linguistics: Essays in honor of Peter Ladefoged*. Orlando, FA: Academic Press, 157–168.

LinearTeam. (2000). The WavNormalizer (Version 1.0.4) [Computer program]. Retrieved 29 Oct. 2007 from http://www.linearteam.dk/.

McDonough, J., Johnson, K. (1997). Tamil liquids: An investigation into the basis of the contrast among five liquids in a dialect of Tamil. *Journal of the International Phonetic Association* 27(1), 1-26.

McGowan, R.S., Nittrouer, S., Manning, C.J. (2004). Development of [ɹ] in young, Midwester, American children. *Journal of the Acoustical Society of America* 115(2), 871-884.

Magnuson, T.J. (2007a). Hikaku onkyou no teddy-yacky: Nihongo no hajikion /r/ to hokubei eigo no kyouseigo hajikion /t, d/ no shokihoukoku [Comparative acoustic teddy-yacky: A preliminary report on Japanese flap /r/ and North American English post-stress flapped /t, d/.] *Proc. 21st Annual Convention of the Phonetic Society of Japan*. 22-23 Sep., 2007, Nagoya Japan.

Magnuson, T.J. (2007b). The story of /r/ in two vocal tracts. *Proc. 16th International Congress of the Phonetic Sciences*. 6-10 Aug., 2007, Saarbrücken, Germany, 1193-1196.

Monnot, M., Freeman, M. (1972). A comparison of Spanish single-tap /r/ with American /t/ and /d/ in post-stress intervocalic position. In: A. Valdman (Ed.) *Papers in Linguistics and Phonetics to the Memory of Pierre Delattre*. The Hague: Mouton, 409-416.

Narayanan, S.S., Alwan, A.A., Haker, K. (1997). Toward articulatory-acoustic models for liquid approximants based on MRI and EPG data. Part I. The laterals. *Journal of the Acoustical Society of America* 101(2), 1064-1077.

Narayanan, S., Byrd, D., Kaun, A. (1999). Geometry, kinematics, and acoustics of Tamil liquid consonants. *Journal of the Acoustical Society of America* 106(4-1), 1993-2007.

Nieto-Castanon, A., Guenther, F.H., Perkell, J.S., Curtin, H.D. (2005). A modeling investigation of articulatory variability and acoustic stability during American English /r/ production. *Journal of the Acoustical Society of America* 117(5), 3196-3212.

Odden, D. (2005). *Introducing Phonology.* New York: Cambridge University Press.

Okada, H. (1999). Japanese. In: *Handbook of the International Phonetic Association: A Guide to the Use of the International Phonetic Alphabet.* Cambridge: Cambridge University Press, 117-119.

Okamoto, M., Ujihara, Y. (1998). *Kiite-oboeru Kansai(Oosaka)-ben Nyuumon* [Listen and Learn Kansai (Osaka) Japanese for Beginners]. Tokyo: ALC press.

Okamura, M. (1995). On'in, onsei [Phonology/phonetics]. In: Kindaiti, H., Hayashi, O., Sibata, T. (Eds.) *Nihongo Hyakka Daijiten* [An Encyclopeadia of the Japanese Language], 231-256.

Onishi, T. (2002). Dialect research at the National Institute for the Japanese Language: Dialectologia et Geolinguistica (DiG). Mouton de Gruyter, 87-96.

Otsuka, N. (1991). Ra/da-gyouon no ayamari ni tsuite no kenkyuu [A study on the confusion between r-gyo and d-gyo sounds]. *Choukaku Gengo Shougai* [The Japanese Journal of Hearing and Language Disorders] 20(2), 69-74.

Palter, D.C., Slotsve, K. (2006). *Colloquial Kansai Japanese: the Dialects and Culture of the Kansai Region.* Tokyo: Tuttle Publishing.

Patterson, D. (2001). Variant frequency in flap production: a corpus analysis of variant frequency in American English flap production. *Phonetica* 58(4), 254-275.

Recasens, D., Pallarès, M.D. (1999). A study of /ɾ/ and /r/ in light of the "DAC" coarticulation model. *Journal of Phonetics* 27, 143-169.

Richmond, S. (2004). Perceived politeness of the Kyoto dialect. *Kyouto Gakuen Daigaku Keizai Gakubu Ronshu* 14(2), 57-72.

Shibatani, M. (1990). *The Languages of Japan.* New York: Cambridge University Press.

Shimizu, K., Dantsuji, M. (1987). A cross-language study on the perception of [r-l]: a preliminary report. *Studia Phonologica* 21, 11-19.

Shingu, I., Hatanaka, J. (2007). Kansai-ben: Kansai Dialect Self-study Site. [Online resource, retrieved 31 Aug. 2008]. MIT, http://llarc.mit.edu/kansai/index.html.

Sourceforge.net. (2006). Audacity: A free digital audio editor (Version 1.2.6) [Computer program]. Retrieved 2 Oct. 2006 from http://audacity.sourceforge.net/ .

Story, B.H., Titze, I.R., Hoffman, E.A. (1996). Vocal tract area functions from magnetic resonance imaging. *Journal of the Acoustical Society of America* 100(1), 537-554.

Sudo, M.M., Kiritani, S., Yoshioka, H. (1982). Nihongo bouin-kan /r/ no chouon: Erekutoroparatogurafii ni yoru kansoku [An electro-palatographic study of Japanese intervocalic /r/]. *Nihon Onkyougakkai Onsei Kenkyuukai Shiryou* [Transactions of the Committee on Speech Research, The Acoustical Society of Japan] S82-30, 233-238.

Sweet, H. (1906). *A Primer of Phonetics*. 3rd Edition. Oxford: Clarendon Press.

Tsujimura, N. (2007). *An Introduction to Japanese Linguistics.* 2nd Ed. Malden, MA: Blackwell.

Tucker, B.V., Warner, N. (2007). Inhibition of processing due to reduction of the American English flap. *Proc. 16th International Congress of the Phonetic Sciences*. 6-10 Aug., 2007, Saarbrücken, Germany, 1949-1952.

UCLA Phonological Segment Inventory Database (UPSID). (1992). UCLA. (Cited from Walsh Dickey 1997).

Vance, T.J. (1987). *An Introduction to Japanese Phonology*. New York: State University of New York Press.

Walsh Dickey, L. (1997). *The Phonology of Liquids*. PhD dissertation, University of Massachusetts Amherst.

Warner, N., Arai, T. (2001). The role of the mora in the timing of spontaneous Japanese speech. *Journal of the Acoustical Society of America* 109(3), 1144-1156.

Warner, N., Jongman, A., Sereno, J., Kemps, R. (2004). Incomplete neutralization and other sub-phonemic durational differences in production and perception: Evidence from Dutch. *Journal of Phonetics* 32, 251-276.

Westbury, J.R., Hashi, M., Lindstrom, M.J. (1998). Differences among speakers in lingual articulation for American English /ɹ/. *Speech Comm.* 26, 203–226.

Wittgenstein, L. (1958). *The Blue and Brown Books. Preliminary Studies for the Philosophical Investigations.* New York: Harper & Row. Cited in Lindau (1985).

Zeng, T. (2007). Understanding flapping in Xiangxiang Chinese: Acoustic and aerodynamic evidence. *Proc. 16th International Congress of the Phonetic Sciences*. 6-10 Aug., 2007, Saarbrücken Germany, 393-396.

Zue, V.W., Laferriere, M. (1979). Acoustic study of medial /t, d/ in American English. *Journal of the Acoustical Society of America* 66(4), 1039-1050.

APPENDIX 1: Accent evaluation task

Accent Evaluation Feedback Form

Thank you for agreeing to participate in this research project. The aim of the research is to describe certain speech sounds used in conversational Japanese, but as the principle researcher is not a native speaker of the language, your help is needed in identifying the variety of Japanese being spoken in the 2 conversations being analyzed.

The 2 Conversations

You will be given two 30-minute digital recordings of conversational Japanese - one between two females ('JFA' [older] & 'JFB' [younger]), and the other between two males ('JMA' [younger] & 'JMB' [older]). For your reference, transcripts of the conversations will also be provided. In evaluating the conversations (see the Questions below), feel free to listen to them as much or as little as you feel necessary.

Questions

1. If you were to guess, how old would you say each speaker is?

JFA: __________ JMA: __________
JFB: __________ JMB: __________

2. What dialect of Japanese would you say each speaker is using?

JFA: __________ JMA: __________
JFB: __________ JMB: __________

3. Compared with a speaker from Tokyo, how 'strong' would you say their accents are? (1 = very slight accent, not different from Tokyo; 5 = very strong accent, different from Tokyo).

JFA: 1 2 3 4 5 JMA: 1 2 3 4 5
JFB: 1 2 3 4 5 JMB: 1 2 3 4 5

4. How formally or informally would you say each speaker is talking?
(1 = very formally; 5 = very informally)

JFA: 1 2 3 4 5 JMA: 1 2 3 4 5
JFB: 1 2 3 4 5 JMB: 1 2 3 4 5

5. In each conversation, is one speaker speaking more formally than the other? If so, who? Could you suggest why?

JFA+JFB?: __________ Why?: ______________________________________

JMA+JMB?: _________ Why?: ______________________________________

That's it - Thanks again for your help!

APPENDIX 2: Summary of drift patterns across all vowel environments, by speaker

JFA (avg., %)	S/N 28.3	M 28.6	W/O 40.6	NOG 2.5	≠RL 36.2	R 30.5	R + L 1.0	L 29.8
/a_a/	○	●	●	●	○	○	○	●
/o_e/	○	○	●	●	○	●	●	○
/a_i/	●	●	○	○	●	○	○	●
/e_ɯ/	●	●	○	○	●	○	○	○
/a_ɯ/	●	○	●	○	○	●	○	○
/a_e/	○	○	●	○	○	○	○	●
/_i/	●	●	○	●	●	○	○	●
/i_/	●	●	○	--	●	○	○	○
/_ɯ/	●	--	○	○	●	●	●	○
/ɯ_/	●	○	●	●	○	●	○	○
/_e/	○	●	●	○	○	○	●	●
/e_/	●	●	○	○	●	○	○	●
/_o/	●	○	●	○	●	●	○	○
/o_/	○	○	●	●	○	●	●	○
/_a/	○	●	●	●	○	●	○	●
/a_/	○	●	●	●	○	○	○	●

JFB (avg., %)	S/N 61.7	M 26.0	W/O 11.0	NOG 1.3	≠RL 74.3	R 6.0	R + L 0.0	L 18.4
/a_a/	○	●	●	●	○	○	--	●
/o_e/	○	○	●	●	○	●	--	○
/a_i/	○	●	●	●	○	○	--	●
/e_ɯ/	○	●	○	○	●	●	--	○
/a_ɯ/	●	●	○	○	●	○	--	○
/a_e/	●	○	●	○	○	○	--	●
/_i/	○	●	○	○	○	○	--	●
/i_/	●	●	○	○	●	●	--	○
/_ɯ/	●	●	○	○	●	○	--	○
/ɯ_/	●	○	○	○	●	●	--	○
/_e/	○	○	●	●	○	●	--	●
/e_/	○	●	●	○	●	●	--	○
/_o/	●	○	○	○	○	●	--	○
/o_/	●	○	●	●	●	●	--	○
/_a/	○	●	●	○	●	●	--	○
/a_/	○	●	○	○	○	○	--	●

JMA	S/N	M	W/O	NOG	≠RL	R	R + L	L
(avg., %)	22.2	24.0	42.3	11.4	6.7	37.6	3.9	40.4
/a_a/	●	○	○	●	●	○	○	●
/o_e/	○	○	●	●	○	●	●	○
/a_i/	●	○	○	○	●	○	○	●
/e_ɯ/	●	●	○	○	●	○	○	●
/a_ɯ/	○	●	●	○	○	●	○	○
/a_e/	●	○	○	●	○	○	●	●
/_i/	●	○	○	○	○	○	○	●
/i_/	○	●	○	●	●	○	○	●
/_ɯ/	○	●	○	○	●	○	○	○
/ɯ_/	○	●	●	○	○	●	○	○
/_e/	○	○	●	●	○	○	●	○
/e_/	○	●	○	○	●	○	○	●
/_o/	○	●	●	○	●	●	○	○
/o_/	○	○	●	●	○	●	●	○
/_a/	○	○	●	●	○	●	○	○
/a_/	●	○	○	●	○	○	○	●

JMB	S/N	M	W/O	NOG	≠RL	R	R + L	L
(avg., %)	29.3	19.1	44.1	7.4	13.8	40.7	4.0	34.0
/a_a/	○	○	●	●	○	●	○	○
/o_e/	●	○	○	●	○	○	●	●
/a_i/	●	●	○	●	--	○	●	●
/e_ɯ/	○	●	○	○	●	○	○	○
/a_ɯ/	●	●	○	○	○	○	○	●
/a_e/	●	○	●	●	○	○	●	●
/_i/	●	●	○	●	●	○	●	●
/i_/	○	●	○	●	●	○	○	○
/_ɯ/	○	●	○	○	●	○	○	○
/ɯ_/	○	●	●	○	●	●	○	○
/_e/	●	○	○	●	○	○	●	●
/e_/	○	●	○	○	●	○	○	○
/_o/	●	○	●	○	●	●	○	○
/o_/	●	○	○	●	○	○	●	●
/_a/	○	○	●	●	○	●	○	○
/a_/	○	○	●	●	○	●	●	●