CHAPTER 2: BACKGROUND AND LITERATURE REVIEW

2.1. Introduction

The palatalization of a velar consonant before a front vowel or glide is one of the most commonly documented sound changes (Bloomfield 1933, Chen 1973, Bhat 1978, Hock 1991). Palatalization of velars has two main outputs; the first and most common outcome is for the velar to become a coronal consonant with a sibilant component, such as [tʃ]; the second result is for the velar to acquire a secondary palatal articulation, as in [kʲ]. Even though velar palatalization has played a role in shaping many languages, there is no consensus on how the sound change happens. The process is usually viewed as articulatorily motivated, but there is little consistency among the various theories and the series of articulatory steps required cannot all be motivated (Sievers 1876, Grammont 1933, Bhat 1978, Hock 1986, Antilla 1989). A perceptual account for the change has been suggested by Ohala (1989). One thing the proposals have in common, however, is the starting point: a fronted velar.

This chapter is organized as follows. First I give examples of velar palatalization from various languages to illustrate its widespread occurrence in diverse language families. Then I review accounts of velar palatalization that offer articulatory and perceptual explanations. I find that the perceptual accounts offer a more simple, straightforward explanation. Following this, I examine the acoustic and perceptual properties of the fronted velar that is postulated in all accounts of palatalization and investigate the acoustic similarities of fronted velars and palatalized velars. Based on the acoustic similarities, I will argue that velar palatalization can be analyzed as a perceptually motivated change and that assuming a perceptual change avoids many of the complications of an articulatorily motivated analysis.

2.2. Examples of Velar Palatalization

In this section, I give examples of the sound change velar palatalization before front vowels and glides. I concentrate on cases in which a velar becomes a coronal affricate, since this subset of velar palatalization is the focus of the experimental work in the following chapters. However, for the sake of completeness, I also give a few examples in which a velar acquires a secondary palatal articulation. By focusing on velar palatalization before font vowels, I do not mean to suggest that this is the only type of palatalization possible or that palatalization is never conditioned by other factors. For instance, Bhat (1978) gives many examples of a dental/alveolar undergoing palatalization (e.g., t > ʃ /__i) and Grimes (1969) and Campbell (1974) give examples of velar palatalization in Mayan which are the result of a dissimilation (k k’ > k^j k^j’ /__V {q q’ x}). Focusing on velar palatalization conditioned by a front vowel or glide simply limits the scope of this work to investigate one of the most common types of palatalization.

I will first consider examples of velar palatalization in which the resultant segment is a coronal. This type of change is conditioned by a palatal glide or front vowel and the resultant segment is often a palatoalveolar affricate. This is especially true for the voiceless velars, but less so for the voiced velars which often become coronal fricatives or palatal glides. In addition, this type of change often results in the creation of a new segment in the language. Following the examples resulting in a coronal, I present a few examples in which the resultant segment is a palatalized velar. This sound change is usually part of a larger innovation in which a language acquires secondary palatalization for most places of articulation.

2.2.1. Velar Palatalization Resulting in a Coronal

In this section, I will provide examples of velar palatalization before a high glide or front vowel resulting in a coronal. This collection is not exhaustive, but it exemplifies the major aspects of this common sound change. The examples here are from a variety of language families, demonstrating the widespread occurrence of this change. A common voicing asymmetry found is that voiceless velars are more likely to palatalize than voiced velars. In Bhat's (1978) collection of examples of velar palatalization conditioned by a following front vowel or glide, about 60% of the cases involved both the voiced and voiceless velars and 40% of the cases involved **only** voiceless velars. There were no examples, however, of a voiced velar palatalizing to the exclusion of the voiceless velar. Also note that in these examples, the voiceless velar is more likely to become a coronal affricate than the voiced velar. For example, in English, Bantu and Slavic the voiceless velar has become a coronal affricate while the voiced velar has become a coronal fricative or palatal glide. In addition, the result of the velar palatalization creates a new segment in all the examples given. In following sections, I will argue that the propensity for voiceless velars to become affricates is due to the acoustic and perceptual quality of their burst and aspiration.

2.2.1.1. Slavic

The Slavic branch of Indo-European provides excellent examples of velar palatalization resulting in a coronal. These sound changes are commonly called the First and Second Slavic Palatalizations. Both happened in the prehistory of Slavic and are a shared development of all the Slavic languages. The presentation and examples for these palatalizations are based on Vaillant (1950), Kiparsky (1963), Shevelov (1965), Matthews (1967), Lunt (1974) and Schenker (1993).

The First Slavic Palatalization changed velars to palatoalveolars before all front vowels and the palatal glide [j], creating new segments. The front Proto-Slavic vowels were: [ĭ], [i] (< PIE *i: or *ei), [e], [ɛ] (< PIE *e:), and [ẽ]. This change happened early in Proto-Slavic. Examples (1-3) show the development of the velars from Pre-Proto-Slavic into Old Church Slavic (OCS), the oldest attested Slavic Language. The OCS forms listed in the "compare" sections offer examples of the morphological alternations upon which a velar reconstruction is based.

(1) k > tʃ / __ {j, ĭ, i, e, ɛ, ẽ}

Pre-Proto-Slavic	OCS	Gloss
*wil**k**-e	vlĭ**tʃ**e	'wolf' (voc.)
*pla:**k**-j-o:-m	pla**tʃ**õ	'I cry'

Compare OCS:

vlĭ**k**ŭ	‘wolf’ (nom. sg.)
pla**k**ati	‘to cry’

(2) g > ʒ / __ {j, ĭ, i, e, ɛ, ẽ}

Pre-Proto-Slavic	OCS	Gloss
*mo**g**-e	mo**ʒ**e	‘was able’ (2-3 sg. aor.)
*lu**g**-j-oː-m	lŭ**ʒ**õ	‘I lie’

Compare OCS:

mo**g**oxŭ	‘was able’ (1 sg. aor.)
lŭ**g**ati	‘to lie’

(3) x[1] > ʃ / __ {j, ĭ, i, e, ɛ, ẽ}

Pre-Proto-Slavic	OCS	Gloss
*dow**x**-e	du**ʃ**e	‘ghost’ (voc.)
*dow**x**-j-oː-m	du**ʃ**õ	‘I blow’

---

[1] The velar fricative is from PIE *s* via the RUKI rule whereby *s* became retroflex after *iː*/*i*, *uː*/*u*, *r*, or *k*. This sound change is shared by the Indo-Iranian and Balto-Slavic branches. The reflex of this sound change is *x* in Proto-Slavic. The velar fricative is also from the PIE aspirated velars k^h and k^{wh}.

Compare OCS:

du**x**ŭ	‘ghost’ (nom. sg.)
du**x**ati	‘to blow’

The Second Slavic palatalization (they are ordered chronologically) also created new segments, the dental affricates. In this sound change, velars became alveolars before secondary [i] and[ɛ] which come from the diphthongs **ai* and **oi*[2]. The velars affected by this change were not changed to palato-alveolars by the first palatalization since the diphthongs had not yet monophthongized and, hence, the conditioning environment was not met. Examples (4-6) show the development of velars into Old Church Slavic.

(4) k > ts / __ {i, ɛ}

Late Proto-Slavic	OCS	Gloss
*vlĭ**k**i	vlĭ**ts**i	‘wolves’ (nom pl.)
***k**ɛna:	**ts**ɛna	‘price’

---

[2] **ai, *oi* > *E* when syllable final word internally and **ai, *oi* > *i* when in the final syllable.

(5) g > dz / __ {i, ɛ}

Late Proto-Slavic	OCS	Gloss
***g**ɛlo	**dz**ɛlo	‘strongly’
*brɛ**g-**ɛ	brɛ**dz**ɛ	‘bank’ (loc. sing)

The reflex of [g] before front vowels is [zʲ] in most of the Slavic languages. We find *zeilo* in Old Czech and *zelo* in Slovenian from **gelo*.

(6) x > s / __ {i, ɛ}

Late Proto-Slavic	OCS	Gloss
du**x**i	du**s**i	‘spirits’ (nom pl.)

The treatment of the velar fricative is subject to dialectal variation. The outcome of the sound change is [s] in East Slavic (Russian, Ukrainian, Belorussian) and South Slavic (Old Church Slavic, Bulgarian, Macedonian, Serbo-Croat, Slovene) and [ʃ] in West Slavic (Czech, Slovak, Upper and Lower Sorbian, Polish, Cassubian). This dialectal split can be illustrated by the treatment of Proto-Slavic **xair-* ‘gray’. Russian (East Slavic) has sɛrŭ, while Old Czech (West Slavic) has *ʃier*ɨ.

The sequences [kw] and [gw] also undergo the second palatalization in the South Slavic and some of the East Slavic languages, but not in West Slavic. Consider the development of the Proto-Slavic words for ‘flower’ and ‘star’ in example (7). The

[kw] and [gw] sequences palatalize in the East Slavic Russian, but not in the West Slavic Polish.[3]

(7)

Proto-Slavic	Russian	Polish	Gloss
kwait-	**tsv**et	**kw**iat	‘flower’
gwaizda:	**zv**ezda	**gw**iazda	‘star’

The Third Slavic Palatalization is sometimes considered part of the second since the outcome is the same, including the dialectal split in the treatment of the velar fricative. In this change, velars were changed into alveolars when they **followed a front vowel**, with or without an intervening nasal. I will not illustrate this sound change since it is outside the scope of this work which is primarily interested in palatalization conditioned by a following front vowel or glide.

To sum up, the Slavic palatalizations offer many examples of velar palatalization before palatal glides and front vowels which result in new consonants at the coronal place of articualation. These examples also exhibit variation in their outcomes. The first palatalization changes [k] to [tʃ], [g] to [ʒ] and [x] to [ʃ], while the second palatalization changed [k] to [ts], [g] to [dz] (and in some languages [dz] > [zʲ]), and [x] to [s] or [ʃ]. Note that the palatalization of the voiceless velar stop produces an affricate in all palatalizations, while the voiced velar stop sometimes becomes an affricate and

---

[3] Note that the [v] in Russian is the reflex of Proto-Slavic [w].

sometimes a fricative. Note also that the result of palatalizing a velar fricative is a dental/palatoalveolar fricative and not a palatal fricative.

2.2.1.2. Indo-Iranian

The Indo-Iranian branch of Indo-European has also undergone a palatalization of velars before front vowels. One of the defining features of the Indo-Iranian branch of Indo-European is what is commonly called the “Law of Palatals”. In this sound change, Proto-Indo-Iranian acquired palatoalveolar affricates through velar palatalization: velars were changed to palatoalveolar affricates in the context of a following front vowel (Mayrhofer 1965 [1972], Hoffmann 1982). Example (8) lays out the sound change and offers examples from Sanskrit and Avestan, which are ancient Indic and Iranian languages respectively.

(8) *k *g *g^h > *tʃ *dʒ *$dʒ^h$ / ___ front V

Pre-Proto-Indo-Iranian	Sanskrit[4]	Avestan	Gloss	c.f. Latin / Greek
***g**i:wo-s (<*g^wihwo)	**dʒ**i:vas	**dʒ**i:va-	'alive'	*vivus* / b…oj
*ju**g**o-m	ju**g**a-	---	'yoke'	iugum / zugÒn
***k**e (<*kwe)	**tʃ**a	-**tʃ**a	'and'	-*que* / te
*se**k**-etoi (<*sek^w-)	sa**tʃ**ate	ha**tʃ**-	'follow'	*sequitur* / ›pomai
***k**ru:-	**k**ru:ra-	**x**ru:-	'bloody, raw flesh'	*cruor* / kršaj

The Law of Palatals thus offers an example in which both the voiced and voiceless velar stops palatalize to palatoalveolar affricates. Also note that the affricates are segments newly created as a result of this palatalization.

---

[4] I have transcribed the Sanskrit using palatoalveolar affricates. It is possible, however, that ancient Sanskrit had actual palatal stops (Allen 1953 [1961]) unlike the modern Indic languages which have palatoalveolar affricates. The traditional analysis holds that the Law of Palatals operated in Proto-Indo-Iranian, and then the palatoalveolar affricates were changed to palatals in Sanskrit. The Avestan reflexes, on the other hand, are agreed to be palatoalveolar affricates. Although, the phonetics of Avestan are hard to retrieve.

2.2.1.3. The Salishan Languages

The Salishan languages can be divided into two groups: the k-languages and the tʃ-languages (Kinkade 1973). Proto-Salish had a plain velar series and a labio-velar series. The k-languages retain the plain velars, while the tʃ-languages have changed all the plain velars to palatoalveolar segments. The tʃ-languages fall into two major geographic regions: (1) Coeur d’Alene and Spokane-Kalispel-Flathead in the Interior Salishan subgroup (the easternmost are the only members of the internal subgroup that are tʃ-languages) (2) All but the Cowlitz language of the Tsamosan Salishan subgroup (note that this does not include Bella Coola, which constitutes its own subgroup). Please see the Figure 2.1 for a subgrouping of the Salishan Languages after Thompson (1979).

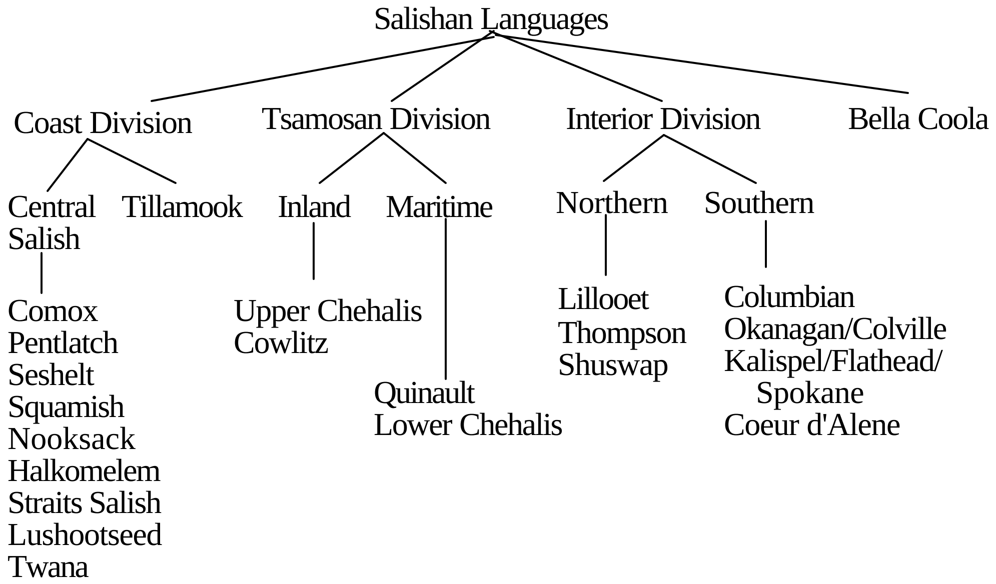

Figure 2.1

Cowlitz Salish has a segmental inventory in between the k-languages and the tʃ-languages. It has both the velar series and the palatoalveolar series. The two places overlap in distribution. Kinkade (1973) offers evidence that the velars in Cowlitz Salish were palatalized before [i] to [tʃ] followed by a phonemic split. Kinkade proposes that this palatalization is the origin of the palatoalveolars in the other Salishan tʃ-languages. In other words, he proposes that velars were first palatalized before [i] as in Cowlitz, then the change was generalized to include all (non-labialized) velars. Note that both the plain and ejective voiceless velar stops become palatoalveolar affricates while the velar fricative becomes a palatoalveolar fricative. This sound change is illustrated in (9).

(9) k, k’, x > tʃ, tʃ’, ʃ /__ i [5]

Proto-Salish	Cowlitz Salish	Gloss
***k’**ilk	**tʃ’**ilk	‘window’
***k**itaq-	**tʃ**æq-	‘argue’
*túl**x**ils-	túl**ʃ**ils-	‘hint’

Thus, the Salishian languages provide another example of velar palatalization in the environment before a front vowel in which a new palatoalveolar segment is created. It must be mentioned that there are no voiced stops in Salishian languages, therefore it is a spurious example of a language in which only voiceless velars undergo palatalization.

---

[5] C’ = an ejective consonant
Note also that the vowel phonemes are i u a ´.

2.2.1.4. Bantu

The Bantu languages show widespread palatalization of velars, which generally become palatoalveolars, before the front vowels **i*, and **e*, as well as through merger with a palatal glide (Guthrie 1967). Hyman and Moxley (1995) have proposed that the Bantu velar palatalization has its roots in merger with a palatal glide and that the velar palatalizations before front vowels are analogical extensions of the palatalization before the glide [j]. For example, they propose that Standard Swahili maintains the original state of affairs. The prefix /ki-/ is realized as [ki] when followed by a consonant (e.g., *ki-jito* ‘brook’) but as [tʃ] when followed by a vowel (e.g. *tʃ-ama* ‘society’). The vowel /i/ of the prefix /ki-/ was glided historically and then merged with the preceding velar to produce [tʃ]. This sound change is reflected in the synchronic alternation of the prefix [ki]~[tʃ].

In other Bantu languages, however, the velar palatalization has been analogically extended to more environments. Mwiini, Kanyok and some dialects of Nyambo palatalize velars in prefixes before front vowels. So the class 7 prefix corresponding to the Swahili prefix above would be [tʃi] (e.g., *tʃi-lòt* ‘dream’). Other Bantu languages have extended the palatalization even further to morpheme initial position (Bemba, Cewa, Tonga and Shona) or to morpheme internal environments (Cifundi and Jomvu), while other languages (Yao, Tumbuka, Mwera, and Nyankore-Kiga) palatalize all velars before front vowels [i] and [e] regardless of morphological environment. (10) gives some examples from Jomvu, which palatalizes velars morpheme internally before front vowels.

(10)

Proto-Bantu	Jomvu	Gloss
*-**k**én**g**édé	**tʃ**en**dʒ**ele	‘bell’
*-po**k**id-	-po**tʃ**e-a	‘receive’
*-in**g**id-	-n**dʒ**i-a	‘enter’

The Proto-Bantu stops were spirantized before [i̊] (a [+ATR] vowel) in most of the Bantu languages (Guthrie 1967, Myers 1992-1994). While the dental and labial places of articulation were also affected by this change, I will only consider the outcome of the velars here. For example, in Shona, the voiced velar stop became a coronal murmured fricative [z̤] and the voiceless velar stop became a coronal affricate [ts] before the high front [+ATR] [i̊].

(11) *g > z̤ / __ i̊

*k > ts / __ i̊

Proto-Bantu	Shona	Gloss
*-**g**ína	**z̤** íná	‘name’
*-**k**índo	mu - **ts**índo	‘audible footstep’

In sum, the Bantu languages offer two examples of velar palatalization. In the first palatalization discussed, a velar stop and a following palatal glide coalesced to form a palatoalveolar affricate. In the second palatalization discussed, a voiced velar became a coronal fricative and a voiceless velar became a coronal affricate before a high front vowel. The second palatalization discussed is similar to the First Slavic Palatalization, in that the voiceless velar becomes an affricate, while the voiced velar becomes a fricative.

2.2.1.5. English

English also offers examples of velar palatalization. By the time of Middle English, the Germanic velar consonants had developed palatal and palatoalveolar sounds under the influence of front vowels. It is generally agreed that the velar consonants before and after front vowels went through a fronted stage in Old English (Emerson 1903, Campbell 1959, Hogg 1979). We find that an Old English fronted [k̟] and [ɣ̟] became [tʃ] and [j] in Middle English word initially before a front vowel (æ ǣ e ē i ī), word internally between two front vowels or between a front vowel and a syllabic consonant, and word finally after front vowels. Examples are found in (12). I do not give examples of velar palatalization conditioned only by a preceding front vowel, since it is not the object of study here.

(12) k̟, ɣ̟ > tʃ, j / {#, front V} __ front V

Old English	Modern English	Gloss
k̟iri**k̟**e	**tʃ**ɚ**tʃ**	‘church’
k̟i:dan	**tʃ**aɪd	‘chide’
k̟e:ake	**tʃ**ik	‘cheek’
ɣ̟inian	**j**ɑn	‘yawn’
ɣ̟eard	**j**ɑrd	‘yard’
ɣ̟eornan	**j**ɚn	‘yearn’

Geminate *k* and *g* also were palatalized word internally under the influence of a front vowel. Fronted [k̟k̟] became [tʃ] and fronted [g̟g̟] became [dʒ]. Examples are found in (13).

(13) k̟k̟, g̟g̟ > tʃ, dʒ / front V __ front V

Old English	Modern English	Gloss
wæ**k̟k̟**ing(e)	wɑ**tʃ**ɪŋ	‘watching’
ky**g̟g̟**el	kʌ**dʒ**əl	‘cudgel’

The exact chronology of the sound changes with respect to *i*-umlaut and ‘breaking’ is somewhat debated (see e.g. Penzl 1947, Kristensson 1976, and Hogg 1979) and many of the changes have been obscured by analogical restoration.

Nevertheless, as can be seen in (12-13), there are many clear examples of the sound change.

The cases of English palatalization discussed here provide more examples of velar palatalization before a front vowel in which the voiced and voiceless segments behave differently. In the first palatalization discussed, the voiced velar became a glide, while the voiceless fricative became an affricate. In the second case, when the voiced consonant was “strengthened” by gemination, it became an affricate. Also note that the affricates came into English *via* velar palatalization.

2.2.1.6. Mam

The dialects of Mam, a Mayan language, also offer examples of velar palatalization resulting in a coronal affricate (England 1990). Most of the dialects change a voiceless velar stop to an coronal affricate before the front vowels [i] and [e]. In one dialect (Tacaná) the result of the palatalization is a palatal stop [c]. In Ostumcalco and Ixtahuacán, the velar is palatalized to a laminal palatoalveolar affricate [tʃ]. In Todos Santos, however, the resulting consonant is an apical post alveolar affricate [tɹ̥].

Examples of the sound change are given in (14). Following the examples are the cognate forms from Quiché (a language in the Quichean branch of the Mayan family). These forms have not undergone the palatalization and are the basis for reconstructing **k* (Cambell 1977).

(14) k > tʃ / __ {i,e} (Ostmucalco and Ixtahaucán)
k > tɹ̥ / __ {i,e} (Todos Santos)
k > c / __ {i,e} (Tacaná)

Ostmucalco	Ixtahaucán	Todos Santos	Tacaná	Gloss	Quiché
tʃe:j	**tʃ**e:j	**tɹ̥**e:j	**c**e:j	‘horse, deer’	**k**e:x
tʃiʔ	**tʃ**iʔ	**tɹ̥**iʔ	**c**iʔ	‘sweet’	**k**i?

The changes in the Mam dialects illustrate the general observation that a coronal affricate is the most common result of palatalizing a voiceless velar. In the case of Mam, three dialects have a coronal affricate as the reflex of [k], while only one has a palatal stop as the reflex of [k].

2.2.1.7. Chinese

Between Old Chinese and Middle Chinese, velars were palatalized before the combination of palatal glide and front vowel . The products of this palatalization were alveolo-palatal affricates and fricatives in Middle Chinese (Baxter 1992, Cheng 1968, Pulleyblank 1984). The alveolo-palatals were allophonic variants of the velars, only found before a palatal glide and front vowel combination. Old Chinese did not have alveolo-palatals (Baxter 1992:177). Therefore, the sound change created new allophonic variants. Example (15) gives the sound change followed by example forms of Old and Middle Chinese. For comparison, example words are also given for which the conditions of the sound change are not met. Examples are taken from Baxter (1992:207 ff.)

(15) k kʰ g x > tɕ tɕʰ dʑ (ʑ) ɕ / __ j front V

Old Chinese	Middle Chinese	Gloss
***k**je	**tɕ**je	‘branch’
***k**jijʔ	**tɕ**jijx	‘fine-tasting (food, wine)
***kʰ**jet	**tɕʰ**jet	‘to trail, drag’
***g**jip	**dʑ**jip or **ʑ**jip	‘ten’
***x**jiw	**ɕ**juw	‘catch, take, collect, receive’
Compare:		
***k**aw	**k**aw	‘tall, high’
***kʰ**aʔ	**kʰ**ux	‘bitter’
***g**aj	**h**a	‘(Yellow) river’
***g**jo/ut	**g**jwot ~**g**jut	‘dig out (earth)’
***x**jat	**x**jot	‘to cease, to rest’

Both the voiceless aspirated and unaspirated velar stops produce affricates. There is some debate however, about whether the voiced stop produced an affricate or a fricative (for discussion see Baxter 1992:52). Note also that both the palatal glide and the front vowel were necessary for this change. As the examples illustrate, a palatal glide by itself was not sufficient.

2.2.2. Velar Palatalization Resulting in a Palatalized Velar

In this section, I will provide examples of the second common output of velar palatalization: a secondary palatalization producing a palatalized velar [k^j]. This outcome is not as frequent as a coronal affricate (Bhat 1978). In addition, the acquisition of secondary palatalization on velars is often part of the development of a new series of palatalized stops for several places of articulation.. This is the case in the two examples discussed here: Russian and Irish Gaelic.

2.2.2.1. Russian

Subsequent to the velar palatalizations of Slavic discussed above, Russian has had many palatalizations resulting in secondarily palatalized consonants. The labials and alveolars acquired palatalized variants first and then the velars later. By the eleventh century, there were separate phonemes for the palatalized and non-palatalized (or velarized) varieties of the bilabial stops and nasals, the dental stops and nasals, the labio-dental and alveolar fricatives, and the liquids (Mathews 1967:153-157). Table 2.1 illustrates the palatalized/non-palatalized series.

Palatalized/Non-Palatalized Distinctions (11th C)

Non-Palatalized			Palatalized		
Bilabial	Labio-dental	Dental	Bilabial	Labio-dental	Dental
p b		t d	p^j b^j		t^j d^j
m		n	m^j		n^j
	f v	s z		f^j v^j	s^j z^j
		l r			l^j r^j

Table 2.1

The secondarily palatalized phonemes are the result of palatalization before front vowels and [j]. The original phoneme acquired a palatalized allophone, and then a subsequent phonemic split created two separate phonemes.

By the thirteenth century, Russian manuscripts begin to show velars written as palatalized before front vowels and [j] and by the fourteenth century, the velars are regularly written as palatalized in that context. At this time, for example, /i/ is regularly written "backwards N" [i] and not "bl" [ɨ] after velars, indicating a palatalized velar consonant.

Today, the Russian velars are undergoing a phonemic split. The non-palatalized velars occur only before non-front vowels, consonants, and word finally, as illustrated in Table 2.2.[6] Palatalized velars occur primarily before front vowels as shown in Table 2.3. However, there are cases where the palatalized velars occur before non-front vowels. These are mainly found in loan words as in Table 2.4 (Jones and Ward 1969).

---

[6] Only voiceless obstruents are found word finally.

Non-Palatalized Velar Before Non-front Vowel

Non-Palatalized Velar	Gloss
ka**k**	‘how’
ku**k**lə	‘doll’
konʲ	‘horse’
ro**k**	‘fate’
got	‘year’
gvalt	‘hubbub’
xatə	‘hut’
mo**x**	‘moss’
ʃa**x**tə	‘mine’

Table 2.2

Palatalized Velar Before Front Vowel

Palatalized Velar	Gloss
kʲit	‘whale’
kʲɛm	‘whom’ (instr.)
no**gʲ**ɪ	‘feet’
gʲenʲɪ(j)	‘genius’
xʲitɨ(j)	‘cunning’
xʲerʲis	‘sherry’

Table 2.3

Palatalized Velar Before Non-front Vowel

Palatalized Velar	Gloss
t**kʲ**ot	‘weaves’
mənʲɪ**kʲ**or	‘manicure’

Table 2.4

2.2.2.2. Irish Gaelic

Irish has a series of secondarily palatalized consonants. Most of the labial, dental and velar phonemes have palatalized counterparts. These palatalized variants were already present by the time of Old Irish (Thurneysen 1946). See Table 2.5 for a list of the consonants having palatalized and non-palatalized variants in Modern Irish.

Palatalized/Non-Palatalized Distinctions in Irish

Non-Palatalized			Palatalized		
Labial	Dental	Velar	Labial	Dental	Velar
p b	t d	k g	pʲ bʲ	tʲ dʲ	kʲ gʲ
m	n	ŋ	mʲ	nʲ	ŋʲ
f w v	s	x ɣ	fʲ	sʲ	xʲ ɣʲ
	r l			rʲ lʲ	

Table 2.5

The palatalized phonemes were created when Pre-Old Irish lost final vowels and syncopated many other vowels. Consonants that were at one time before a front vowel (*i* or *e*) acquired a phonological specification for a palatalized articulation. Case marking is often accomplished by the presence or absence of a palatalized consonant whereas it was previously marked by a suffix of different vowel type. Thus, palatalized consonants play a large part in Irish morphology. For example, in Modern Irish the nominative singular for 'man' is *fear* [fʲær], while the genitive singular is *fir* [fʲirʲ]. In the prehistory of Irish, the nominative suffix for this declension was -*as*, and the genitive was -*i*. When final vowels were lost, the palatalizing effect of the front vowel was phonologized as part of the final consonant ([rʲ] in this case) and a new phoneme was created.

Irish, like Russian, provides an example of palatalized velars originating before front vowels in a language with widespread secondary palatalization. Note also that the conditioning environment was sometimes lost in Irish. In Russian, the conditioning environment was lost in the case of a palatal glide but not a following front vowel.

2.3. Accounts of Velar Palatalization as a Sound Change

Traditionally, the sound change of palatalization has been conceived as a gradual articulatory change. There is not, however, a consensus view on how the articulatory change is accomplished. Much of the traditional literature does not offer a detailed description of palatalization. Those authors who do offer a description vary considerably in their accounts. I suggest that the difficulty these accounts face is due to the assumption that palatalization is purely an articulatorily motivated change. Most accounts begin with a fronted velar, but diverge from there. While it might be possible to imagine how the tongue begins at a velar articulation and gradually creeps up to a palatoalveolar place of articulation, changing the part of the tongue used for the articulation (from dorsum to blade) as well as acquiring a fricative release along the way, we will see below that when one actually tries to flesh out an articulatory account, it is problematic. In this section, I review some articulatory accounts of velar palatalization and note that many accounts fail to adequately motivate a change from a palatal to a post-alveolar place of articulation as well as a change from dorsal to laminal/apical articulation. Following this, I review a perceptual alternative proposed by Ohala.

Let us first consider the Neogrammarians, who thought that sound change happened through gradual changes in articulation. Paul (1937) describes the process of gradual articulatory change in which the ‘motor sensation’ or *Bewegungsgefühl* is gradually altered. The *Bewegungsgefühl* consists of a set of articulations for a particular sound. An analogy often given for the *Bewegungsgefühl* is that of an archer aiming at a target. It is rare that the bull’s-eye is hit, but, in general, the arrows are evenly distributed around the bull’s-eye. This set of articulations constituting the *Bewegungsgefühl* is continually being redefined as older articulations are forgotten and newer ones take their place. In the archer analogy, this would be equivalent to the target following the arrows. If the archer landed more arrows to the right of center, the target would move to the right. In this way a sound can change if the *Bewegungsgefühl* is shifted by a preponderance of articulations in one direction. Thus, if a speaker often produces a given segment with articulations that are to one side of a ‘target’, the segment may move in the direction of the frequently produced articulation. This shift may be precipitated by the effect of surrounding segments or by considerations of energy expenditure. Paul is careful to mention that ‘convenience’ is a secondary cause of sound change and that *Bewegungsgefühl* is the primary cause. It should be noted that the Neogrammarians thought the the ‘sound-picture’ or *Lautbild* (a notion of how a word or segment should sound based on what the speaker has heard) constrained the rate at which the *Bewegungsgefühl* could be altered. Thus, the change in the *Bewegungsgefühl* is counterbalanced by the *Lautbild* so that an individual speaker can only change their *Bewegungsgefühl* more or less as quickly as the the other members in their speech community do (as reflected by the *Lautbild*).

Palatalization is considered a type of assimilation by the Neogrammarians. Thus, [k] was thought to move forward by gradual changes in the *Bewegungsgefühl* conditioned by the following vowel. Thus, the first step in the sound change was necessarily a fronted [k]. These movements were imperceptible as they were balanced by the *Lautbild*. Sievers (1876) describes the palatalization of [k] before a front vowel rather vaguely as an equalization of the tongue articulation:

> Beispeile für die Berührung von Consonanten mit Vocalen sind der Eintritt der Mouillirung and Labialisirung, soweit diese auf Ausgleichung der Zungenarticulation beruhen; also namentlich die Verlegung der Articulationsstellen der *k*-Laute je nach dem folgenden (seltner dem vorhergehenden) Vocale, z.B. ihre Palatalisirung vor *e*, *i*, *ö*, *ü*.
> [Examples for the contact of consonants and vowels are the indroduction of softening and labializing [of the consonant], as far as they rely on the equalization of the lingual articulation. Thus, the transfer of place of articulation of a [k] before a following vowel (or more rarely a preceding vowel), for example the palatalization of [k] before *e*, *i*, *ö*, *ü*.] (Sievers 1876:138) (This and all following translations are mine.)

Grammont (1933) proposes a more detailed account of the gradual articulatory change for the palatalization of [k] before front vowels resulting in [tʃ]. The first step in this change was for the velar to be fronted to the hard palate. At this point the fronted [k] could be phonologized as a palatalized [k^j] or develop a voiced fricative release. If a fricative release was developed, the next step would be for the articulation to be moved to a more anterior dental/alveolar place which would still be articulated with the body of the tongue. Grammont offers no explanation for the connection between frication and fronting, despite an account that essentially links the two. He offers no reason why the palatalization cannot stop at an affricated palatal. He goes on to posit what could happen to the alveolar affricate (articulated with the body of the tongue and having a voiced release). If the articulation were fortis, the fricative release would devoice and become a voiceless fricative of the language in question such as a [s] or [ʃ] the result of which would be the affricate [ts] or [tʃ].

> Il est fréquent que les voyelles antérieures *i*, *é*, *œ* fermé, *ü*, *a* antérieur, attirent à elles un *k* qui les précède, dont le point d'articulation passe alors en avant du sommet de la voûte palatine. Dans cette position l'explosion est facilement suive d'un élément fricatif du genre *y* ; si cet élément ne se développe pas, le résultat est un *k* mouillé, *k'*: alb. *k'int* de *centu* ; s'il se développe le résultat est un *ky*, et, pour peu que l'ariticulation avance encore, l'explosion n'a plus lieu dans le domaine du *k*, mais dans celui du *t*, d'où *ty*, avec un *t* articulé la pointe appuyé contre les alveoles des incisives inférieures. Si l'articulation est molle, par example devant l'accent, *ty* subsise: Vionnaz *tyevra* de *capra* ; si l'ariticulation de *t* est violente, par example après l'accent ou à l'initial devant voyelle inaccentuée, le *y* s'assourdit. La plupart des langues n'ayant pas de *y* sourd, le remplacent part ce qu'elles ont de plus voisin, *s*, *s'*, ou *ˇs*, pour lesquels la langue a déjà à peu près la position requise, avec se pointe en bas et une gouttière plus ou moins nette sur sa partie antérieure ; *ts* suppose une articulaiton plus tendue que *tš*, avec les mâchoires un peu plus écartées; pour *tš* la lange est étalée plus largement et plus mollement sur la voûte palatine: Vionnaz *tse̥vo* de *caballu*, *letse* de *leccat*..
> [Frequently the front vowels [i é œü æ] attract a preceding [k] whose place of articulation moves forward anterior to the palatal arch. In this position, the release is easily followed by a fricative component such as

> [ ]. If this component does not develop, the result is a palatalized *k*, [kʲ] (e.g. the Albanian *k'int* from *centu*). If the fricative component develops, the result is a [kj̝]. The articulation only has to move a bit further forward to be in the domain of [t]; here the stop portion of [tj̝] is articulated by pressing the tongue against the socket of the lower incisors. If the articulation is lenis, for example before the accent, [tj̝] remains (e.g. Vionnaz*tyevra* from *capra*). If the articulation is fortis, for example after the accent or word initially before an unaccented vowel, the [j̝] is devoiced. Most languages do not have a voiceless [j̝]; it is replaced by whatever is closest in the language : [s], [sʲ] or [ʃ] for which the tongue already has more or less the correct position with the tip lowered and a trough in the anterior portion. [ts] has a more tense articulation than [tʃ], with the jaw slightly more open; for [tʃ], the tongue is more spread out over a greater area and less rigid (e.g. Vionnaz *tse̥vo* from *caballu*, *letse* from *leccat*).] (Grammont 1933:214)

Hock (1991) also has a detailed description of how a velar palatalization before a front vowel takes place. He proposes that a velar first develops an offglide. The process may stop at this point[7] or continue on. If the palatalization is continued, the next step is for the palatalized velar to become a palatal consonant. Hock describes this process as phonologically motivated but notes the acoustic and articulatory similarity of a palatalized velars and palatals.

---

[7] Note that this first step is the only one that Hock considers to be true palatalization. The following developments, according to Hock, obscure the effects of palatalization and are often confused with palatalization.

> [T]he motivation for this development seems to be phonological, and not just purely phonetic: As palatalization becomes unpredictable and as consequently palatalized and nonpalatalized segments come to be in contrast, there is a tendency to mark this new phonological distinctiveness by an increase in phonetic distinctiveness. This process has been called **polarization**. For velars…such polarization is accomplished most easily by shifting of the palatalized segment toward the palatal area of articulation: Articulatorily, they are similar to palatals…This is matched by a similar acoustic affinity between palatals and palatalized velars and dentals. (Hock 1991:75-76, bold in original)

After the change from velar to palatal, the palatal segment which is "characterized by an inherent sibilant-offglide" (Hock 1991:76) can be assibilated. In other words, the palatal [c] may acquire a fricative release. The fricative is articulated, according to Hock, either with the body of the tongue on the palate [cç] or with the tip of the tongue in a post-dental position [cʃ]. The segment [cʃ] can then change to [tʃ] as the tip of the tongue comes to be used at a place behind the alveolar ridge for the stop portion as well. It is then possible to lose either the stop or fricative portion of the affricate and end up with [t], [ʃ] or [s]. Thus, Hock describes the change from palatal to post-alveolar as the result of (1) changing the articulation of the fricative from palatal to post-alveolar and (2) moving the place of the stop occlusion in assimilation to the fricative. He does not, however, provide an explanation for the shift in the place of articulation of the fricative from palatal to post-alveolar.

Anttila (1989) also offers an explanation of velar palatalization. He proposes that a fronted [k] can be shifted to a coronal place of articulation due to the narrow shape of the vocal tract, producing a palatalized [t^j]. The palatalized [t^j] is in turn assibilated to a [tʃ] or [ts]. This assibilation, according to Anttila is due to a voicing assimilation of the offglide to the stop and an assimilation of tongue shape to the following high front vowel.

> [Palatalization] is an assimilation of a consonant to a high front vowel, usually following; it is an assimilation of tongue position. The vocal tract is narrow for the front vowel, and the stop is shifted into or toward this area, as in English [ḱi:p] vs. [ku:l] (or [ḱiyp] vs. [kuwl]) *keep* vs. *cool*. A palatalized [ḱ] can shift further front and give [t́]. When this affricates, the result of [tʃ] or [ts]. (Anttila 1989:72)

Anttila describes the palatalization of Latin [k] to French [s] (e.g. centum > cent [sɑ̃]; cinis > cendre [sɑ̃ːdʁ])as follows.

> [I]n reality *k* palatalized first into ḱ, and this gave t́ = t^y. then the narrow transition from stop to vowel [y] was assimilated to the voicelessness of the stop part and the groove tongue shape of the following vowel, giving *ts*. The assibilation was completed when closure disappeared altogether, yielding *s*. (Anttila 1989:73)

Thus, Anttila offers an articulatory explanation for the change in place of articulation from palatal to alveolar: the narrow shape of the vocal tract. He does not, however, explain why a narrowed vocal tract would give rise to fronting.

Bhat (1978) examines about 120 cases of palatalization from many different language families and offers a typology distinguishing three types of palatalization: (a) tongue fronting, (b) tongue raising and (c) spirantization. He supports the claim that these three processes are independent by citing examples where they occur alone as well as in combination with each other. Tongue fronting (a) is found in the palatalization of velars where the place of articulation is moved forward: either more anterior on the soft palate, or onto the hard palate. Bhat offers evidence that velar fronting is most often found in stressed syllables before front vowels. Tongue raising (b) can be found in the palatalization of dentals/alveolars, labials and velars. Apical consonants typically become laminal and are moved slightly back. Most often, dentals/alveolars become palato-alveolars. These apical consonants are usually palatalized in the environment of high front vowels or palatal glides. The effect of tongue raising on velars creates alveolars or palato-alveolars. Bhat also considers the addition of a secondary articulation in the form of a palatal offglide a type of tongue raising. This secondary articulation can be added to labials, apicals and velars. Bhat notes, however, that a secondary palatal articulation is comparatively less frequent than shifts in place of articulation (p.67). Spirantization (c) is defined as the addition of frication or stridency to a consonant. This process affects velars, apicals, and palatals, but rarely the labials.

The most common palatalization for each place of articulation are also listed. Velars usually undergo fronting, raising, and spirantization before stressed front vowels, thus producing a change from a velar to a palatoalveolar affricate. Apical consonants usually undergo spirantization before high front vowels or a palatal glide, thus producing a change from a dental/alveolar to an alveolar or palatoalveolar affricate. Labial sounds most often add a secondary palatal articulation. The palatal glide can also undergo palatalization in the form of spirantization.

It is generally agreed that the most common palatalization is that of a velar before a front vowel or glide. Hock (1986), Bhat (1978), Chen (1973) and Bloomfield (1933) all write that velars are one of the the most commonly palatalized segments.

Hock (1986), Chen (1973) and Neeld (1973) also propose an implicational hierarchy for conditioning environments of palatalization. The most preferred is the palatal glide [j], then the front vowel [i] followed by [e] and [æ] respectively. Chen writes:

> Thus even though we cannot predict that palatalization will take place in language X, we can nevertheless predict that if palatalization occurs at all, it will spread along two dimensions of axes: (i) palatalization spreads from the back to the front consonantal series, that is the first consonant series to undergo palatalization will be the velar series, the next one will be the dental, and the last will be the labial series; (ii) palatalization of a prevocalic consonant spreads from high to low front vowels: if [k] palatalizes before a low front [a], so it must also before a mid [e] and a high [i]. (Chen 1973:177)

Ohala (1994) explains the hierarchy proposed by Chen and Neeld on perceptual grounds. He writes that a consonant is more likely to develop a palatal transition into a front vowel the higher that vowel is. If listeners then dissociate the transition from the surrounding segments, the consonant is more likely to develop an offglide before higher front vowels than lower since there is more likely to be an offglide before these segments. Ohala (1989) has also suggested a perceptual motivation for the palatalization of [k] before front vowels. He mentions the high confusion rate of [ki] for [ti] in a perceptual study by Winitz et al. (1972) and suggests that the sound change [k] > [tʃ] can occur if the listener fails to factor out the effects of [i] on the velar.

I will investigate Ohala’s proposal that palatalization is a perceptually motivated sound change. I suggest that the reason traditional explanations fail to motivate each step in the palatalization process on purely articulatory grounds is due to the fact that important aspects of velar palatalization are perceptually motivated. This does not deny articulation a role in the sound change. Indeed, I agree with the proponents of articulatory explanations that the origin of the sound change is to be found in fronted velars. However, as noted above, the problem with many of the articulatory accounts is that they fail to explain why a palatal segment would move forward to a post-alveolar position and change from a dorsal to an apical/laminal articulation. I suggest that the explanation for this is to be found in the acoustic/perceptual properties of a fronted [k]. I begin outlining a proposal for the perceptual conditioning of velar palatalization by reviewing acoustic and perceptual studies which make some note of velars before front vowels. I then review perceptual and acoustic studies of coronal affricates and fricatives and palatalized stops with an eye toward finding similarities between a fronted velar and [tʃ].

2.4. Acoustic and Perceptual Properties of Fronted Velars

It has long been known that a velar is fronted before front vowels. Articulatory studies confirm this notion. For example, Jones (1922) notes that a [k] pronounced before a front vowel makes an imprint on a palatogram, while a [k] pronounced before a non-front vowel makes no imprint on a palatogram. (The palatogram does not extend to the soft palate.) Keating (1993) and Recasens (1990) also review many x-ray and palatographic studies which illustrate the fronted nature of a velar before a front vowel. In general, a fronted velar is produced on the postpalate or rear of the hard palate whereas a non-fronted velar is produced on the soft palate. The difference between the articulations is somewhere in the 5-10mm range. The fronted varieties do not overlap with the non-fronted velars in their area of articulation. Thus, the two types of velars really have separate places of articulation.

In the following sections, I focus on the acoustic and perceptual properties of fronted velars. I review studies dealing with acoustic cues to place of articulation in stops looking for characteristics of velar stops similar to affricates. I also review perceptual studies which include velars and point out cases in which a velar has been confused with a coronal.

2.4.1. Acoustic Cues to Stop Place of Articulation with a Focus on Fronted Velars

Many papers have been written on the acoustic cues to stop place of articulation. Here I will review these papers with an eye toward the velars. One thing all the papers have in common is the finding that the velars behave differently from the other places of articulation in that they show a greater acoustic difference before different vowel types. In other words, the coarticulatory effect that front vowels have on velars is strongly represented in the acoustic signal–the velars are literally split into two classes.

Before continuing with the literature review, I will clarify some of the terminology. Throughout the discussion, I use Fant's (1968, 1973) terminology and division of a stop articulation. Fant decomposes a fully developed stop into four or five successive segments. The first of these is the **occlusion**. There may be voicing or not during this segment. The second segment is the **transient** which is the response of the vocal tract to the pressure release when the occlusion is released; it excludes any turbulence effects. The duration of the transient can be from 2 to 30 ms. long but generally is less than 10 ms. The next segment is the **fricative**; it is characterized by noise produced at the consonantal constriction in the same way as a homorganic fricative. Both the transient and the fricative segments are subject to zeros which act to cancel back cavity formants while front cavity formants prevail. The **aspirative** segment is next. This is not generally found in ‘voiced’ stops. This segment is the product of noise originating from a random source at the glottis or from a supraglottal source with a relatively wide constriction. All vocal tract formants are excited by this noise. Finally, the last segment is composed of the **transitions** of the following vowel formants to the extent that they are influenced by coarticulation with the stop.

The term **burst** is also used in describing stop segments. Sometimes it is used to mean the transient and fricative segments combined and sometimes it is used to describe all the aperiodic segments: the transient, the fricative and the aspiration segment. In the following sections, I will use the term burst to refer to the transient and fricative portions combined since they are hard to separate in natural speech and are analyzed as a single unit in many acoustic studies of stop consonants. I will, however, refer to the aspiration as separate from the burst.

Many studies of acoustic cues to stop place of articulation have been performed since the first codification of spectrographic cues to place and manner of articulation found in *Visible Speech* (Potter *et al.* 1947). In general, it is agreed that both the burst and the formant transitions provide information about place of articulation. Some studies have focused on either the burst or the transitions, while some have considered both. For a clear presentation, I will first consider information from the burst and aspiration, then from the transitions.

2.4.1.1. Burst and Aspiration

Descriptions of the acoustic characteristics of stop burst spectra fall into two main categories: those that investigate static characteristics of the burst, and those that consider dynamic characteristics of the burst. The investigations into static characteristics look at the spectral peaks or general spectral shape of the first 10 to 30 ms. of the burst. The investigations into dynamic properties of the burst look at duration and/or changes in spectral peaks over the length of the burst. Yet another property of the burst to be commonly investigated is the overall duration. Burst duration is often measured with aspiration.

The studies investigating the static characteristics of the burst generally find that velars have one predominant spectral peak whereas the labials and alveolars have a spectrum in which the energy is more evenly distributed. The prominent velar spectral peak is also found to vary in the frequency domain from 1000 to 4000 Hz as a result of vowel context. Studies focusing on the dynamic spectral properties of stop burst and aspiration have found that velars have a later voice onset than other places of articulation and that velars have a mid-frequency peak (1000 to 4000 Hz) which lasts for 15 ms. or longer. Studies focusing on the duration of burst and aspiration have found that velars have longer transients, frication portions, and aspiration than the other places of articulation. In general, stops also have a longer voice onset time and more frication before high vowels. Thus, a velar before a high vowel would have the most frication and longest aspiration period. A velar before a high vowel would sound more like a fricative than other stop productions. For our purposes, it is also interesting to note that velars before high front vowels were confused with alveolars in studies that visually evaluated burst spectra.

Let us first review studies that investigate the static cues to place of articulation in stop bursts. These studies find that velar segments have two distinct patterns, one before the unrounded front vowels, and one before the other vowels. The other places investigated (labial and dental/alveolar) in general only have one pattern.

Jakobson, Fant, and Halle (1952), Stevens and Blumstein (1978, 1981), and Blumstein and Stevens (1979, 1980) all consider gross spectral characteristics of the stop burst for the labial, alveolar and velar place of articulation in terms of the features compact and diffuse. According to this distinction, compact phonemes are characterized by a relative predominance of one centrally located formant region, whereas in diffuse phonemes one or more non-central formant regions predominate. Velars are considered compact and labials and dentals diffuse.

Fischer-Jørgensen (1954), Halle *et al.* (1957), Fant (1973), and Zue (1976) have examined the acoustic patterns of stop bursts before a variety of vowels and found that the compact velar burst spectrum varies as a function of the following vowel. Fant (1973) reports on the burst characteristics of Swedish stops /p b t d k g/ before nine Swedish vowels. Spectrographic investigation revealed that /p/, /b/, /t/, and /d/ had spread-out spectral energy, with the energy located lower for the bilabials than for the alveolars. For /g/ and /k/, the peak frequency of the first 10-30 ms. of the burst was correlated with F3 and F4 before the mid and high front unrounded vowels, with F3 before mid and high front rounded vowels, and with F2 before back rounded and unrounded vowels.

Halle *et al.* (1957) measured the first 20 ms. of the burst for English /p b t d k g/ before the vowels /i ɪ ʌ ɑ u/. They found that the primary concentration of energy for /p/ and /b/ was from 500 to 1500 Hz. /t/ and /d/ had a flat spectrum or one in which the higher frequencies (above 4000 Hz) predominated. /k/ and /g/ were found to have a spectral peak between 2000 and 4000 Hz before front vowels, while the peak was lower before back vowels.

One of the first acoustic studies on stops was performed by Fischer-Jørgensen (1954). She did a spectrographic analysis of the peak frequencies of the burst in the Danish stops /p b t d k g/ before all the Danish vowels. Unlike the other studies, she found that the velars /k/ and /g/ had two maxima. The higher was constant around 4000-5400 Hz. The lower maxima varied with the following vowel: for the round vowels (front and back) the lower maxima followed F2, and for the front unrounded vowels the maxima was closer to F3. Note that the formant correlation found by Fischer-Jørgensen for Danish velars is different from that found for Swedish by Fant. She also found that /t/ and /d/ have maxima at 1950, 3000, 3800, and 4800 Hz. Before a rounded vowel, the lowest two are stronger, and before an unrounded vowel the higher two are stronger. /p/ and /b/ have weaker bursts with energy evenly distributed among the peaks at 1200-1300, 1800, 2700, 3800 and 4800 Hz.

Zue (1976) measured the peak spectral frequency of the burst for English /t d k g/ before the vowels /i ɪ eɪ ɛ æ ʌ ɑ o ɥ u ʊ ɚ aɪ ɔɪ aʊ/. The measurements were taken from LPC spectra 10-15 ms. after the release of the consonant. Both the velars and alveolars were found to have clear peak frequencies. Zue notes, however, that the spectra for the labial stops had no clear peak frequencies and, therefore, he did not report the results. The velars showed variation in the peak spectral frequency as a function of the following vowel. The average peak spectral frequency for a velar before a front vowel was 2720 Hz, before a back unrounded vowel was 1770 Hz, and before a back rounded vowel was 1250. The alveolars showed less variation according to vowel context. Alveolars before rounded and retroflexed vowels had an average peak spectral frequency of 2950 Hz and alveolars before all other vowels had a peak spectral frequency of 3530 Hz.

Based on perceptual test with synthetic burst stimuli (see §2.4.2. for a discussion of perceptual tests) Stevens and Blumstein (1978) came up with a set of spectral templates for the burst of labial, alveolar and velar stops. Labials had a diffuse falling spectrum. Alveolars had a diffuse rising spectrum. Velars had a compact spectrum with a prominent spectral peak in the mid-frequency range from 1-3 kHz: the peak is between 1 and 2 kHz before /u/ and /a/ and around 3 kHz before /i/.

These spectral templates were then matched to stop bursts in human speech. Blumstein and Stevens (1979) made LPC spectra of the first 26 ms. of the burst in CV utterances beginning with /p b t d k g/ followed by a variety of English vowels. These spectra included aspiration in the case of voiceless stops, and possibly part of the vowel in the voiced stops. The spectral templates were fit to the LCP spectra at an 85% correct rate. Note, however, that velar consonants tended to be misclassified as alveolar in the context of high front vowels.

Now let us move on to dynamic burst cues to place of articulation. Several researchers have investigated spectral changes over time in the burst spectrum (Searle *et al.* 1979, Kewley-Port 1983, Kewley-Port *et al.* 1983, Lahiri *et al.* 1984, Suomi 1985, and Hawkins and Stevens 1987). Kewley-Port (1983) had subjects **visually** evaluate burst spectra for a CV beginning with /b d g/ and followed by a variety of vowels. For each CV, there were 8 spectra taken at 5 ms. intervals. There were two types of spectral analysis used: LPC and, following Searle et. al. (1979), spectral representations which model the auditory spectra more closely. The resulting display illustrated the changing spectrum over the first 40 ms. of the CV. Note that these spectra often went into the vowel. The subjects were to judge the 'running' spectra for three features. The first was *Tilt of the Spectrum at Burst Onset*. A labial should be judged to have a flat or falling first frame while an alveolar should be judged to have a rising first frame. The second feature was *Late Onset of Low-frequency Energy*. A velar should be judged to have late onset of low frequency energy starting in the fourth frame or later. The last feature was *Mid-frequency Peaks Extending Over Time.* A velar should be judged to have a single prominent peak between 1000 and 3500 Hz occurring for three or more frames. The subjects were able to assign the running LPC spectra to the correct consonant (/b/, /d/, or /g/) 88% of the time. Note that /g/ was poorly identified before /i/, being identified as a /d/. Results varied for the different auditory spectral representations used.

To summarize thus far, studies investigating the acoustic properties of stop bursts have found that velars have a compact spectrum, the amplitude peak of which varies in the frequency domain as a function of the following vowel. The labial and alveolar/dental stops, on the other hand, have a more diffuse spectrum which is not as variable across vowel contexts. There is also evidence from visual matching of spectral templates to spectra that velars before a high, front, unrounded vowel are confusable with alveolars.

Studies focusing on the duration of burst and aspiration (if present) of stops have found that velars have longer durations and that stops before high vowels have a longer burst and aspiration periods. Thus, a velar before a high vowel will have the longest duration of burst and aspiration, resulting in the velar production closest to an affricate. First I briefly review investigations which study the effect of consonantal place on the length of stop burst and aspiration, then I discuss the findings which indicate that a following high vowel induces a longer burst and aspiration period in a preceding stop.

Lisker and Abrahamson (1964), Klatt (1975), Tekieli and Cullinan (1979), and Krull (1991) have all noted that the burst and aspiration (if present) is longer in velar stops than in alveolar or labial stops. Lisker and Abramson (1964) studied voice onset time (VOT) for stops in 11 languages (Dutch, Puerto Rican Spanish, Hungarian, Tamil, Cantonese, English, Eastern Armenian, Thai, Korean, Hindi and Marathi). They measured VOT from the release of the stop to the onset of voicing. They found that velars had a consistently longer VOT for the voiceless series in all languages, whether or not the stop was ‘aspirated’. They also looked at VOT in two speech styles in the 11 languages: (1) the subject read the words in isolation and (2) the subject made up sentences for the target word in initial and medial position. They found that VOT was not necessarily shorter in the sentence condition than in the citation condition for most of the languages. However, in the case of English the VOT for sentences was shorter overall.

Klatt (1975) investigated the burst and aspiration in English word initial stops and consonant clusters by spectrographic measures. He found that word initial /g/ has a longer burst than /d/ or /b/ and that word initial /k/ had a longer burst than /t/ or /p/ (though the burst was very faint and sometimes non-existent in /p/). The aspiration portion of /k/ was not necessarily longer than the aspiration portion of /t/. The total VOT was, however, longer for /k/ than /t/ or /p/.

Tekieli and Cullinan (1979) measured the length of the aperiodic noise (i.e. the transient, frication and aspiration if present) in word initial English stops and affricates before a variety of vowels. They found the following mean durations for aperiodic noise: /b/ 10 ms., /d/ 12 ms., /g/ 22 ms., /dʒ/ 45 ms., /p/ 53 ms., /t/ 59 ms., /k/ 86 ms., and /tʃ/ 86 ms. Again, the velars have a longer VOT than the alveolars or labials. Notice also that /k/ and /tʃ/ have the same mean duration for aperiodic noise, revealing one more dimension in which we find similarity between the two segments.

Krull (1991) reports on her study of word initial VOT in Swedish /p t k/ for spontaneous and citation speech. She found that in spontaneous speech, as well as in citation forms, the velar stop had a longer mean VOT than the dentals and labials. The mean VOT for the dental consonant was in most cases longer than that of the labial. She also found that there was no significant difference between the VOTs for citation and spontaneous speech. There was, however, a large difference both in the duration of the stop gap and the duration of the following vowel: both were longer in citation form.

Fischer-Jørgensen (1954) and Fant (1973) have also noticed that velars have longer transients than other stops and that velars can have a multiple transients. Fant suggests that these multiple transients could be due to a suction reaction at the articulatory stricture caused by the Bernoulli pressure just as in the normal voice source (i.e. the glottis).

Effects of the following vowel on the duration of the burst and aspiration have been studied by Klatt (1975), Tekieli and Cullinan (1979), and Ohala (1983). They have found that the the length of frication and aspiration is a function of the following vowel. There is longer frication and aspiration before a high vowel. The greater frication is a result of the fact that the high velocity airflow created upon release of a stop lasts longer when a stop precedes a high vowel as opposed to a low vowel. In general, aspiration is longer before a high vowel since, at the release of the stop, the time required to reduce the oral pressure to a level conducive to voicing (voicing needs a transglottal pressure drop) takes longer for the smaller cavity of a high vowel than the larger cavity of a low vowel.

In sum, velars have longer (possibly multiple) transients, longer frication, and longer aspiration (if present) than other places of articulation. Stops also have longer aspiration and frication before high front vowels. Thus, a velar before a high vowel should have the longest frication and aspiration period of all the stops. I am proposing that velar palatalization before front vowels is a perceptually conditioned reanalysis of the long frication and aspiration of a velar before a front vowel, resulting in an affricate such as [tʃ]. Given the finding that velars have the most frication before high vowels, we might expect to find affrication of velars before high back vowels as well. The palatalization of a velar before a high back vowel is not, however, a common occurrence. The answer to this apparent problem lies in the spectral properties of the velar release. As mentioned earlier, a velar before a front vowel has a higher peak frequency than a velar before a back vowel. I present evidence below that the velar release before a high front vowel is acoustically similar to a coronal fricative such as [ʃ].

2.4.1.2. Formant Transitions

Vowel transitions have also been investigated for cues to stop place of articulation. Most studies which have investigated cues to stop place of articulation in vowel transitions have focused on word initial stops. By and large, these studies have been concerned with finding an invariant cue to place of articulation. The studies discussed here take one of two tacks. One is to look for the virtual starting point of a formant transition for a given place of articulation. The other looks for an invariant relation between the actual starting point of a vowel and the midpoint of the vowel for a given place of articulation. The terminology used by these two approaches is somewhat confusing as the same word, **locus**, is used to mean something different in each of the frameworks. A locus is the virtual starting point of the formant transition for the first approach and the actual starting point of the formant transition in the other approach. These studies have in common the finding that velars have two distinct formant patterns reflected in the following vowel: one for the front vowels and one for the other vowels.

Following, Potter *et al.* (1947), many researchers (Cooper *et al.* 1952, Fischer-Jørgensen 1954, Liberman *et al.* 1954, and Delattre *et al.* 1955) have looked for a virtual starting point of a formant transition for a given place of articulation which is insensitive to the following vowel. In general, these studies have been concerned with the F2 transition. Delattre *et al.* (1955) sum up this endeavor as follows:

> If we further assume that the relation between articulation and sound is not too complex, we should suppose … that the second-formant transitions rather directly represent the articulatory movements *from* the place of production of the consonant *to* the position for the following vowel. Since the articulatory place of production of each consonant is, for the most part, fixed, we might expect to find that there is correspondingly a fixed frequency position—or "locus"—for its second formant; we could then rather simply describe the various second-formant transitions as movements from this acoustic locus to the steady-state level of the vowel, wherever that might be. (Delattre *et al.* 1955:769)

With the exception of Fischer-Jørgensen, who performed an acoustic analysis of Danish CV utterances, these studies investigate the perception of synthetic stimuli. F2 transitions are systematically varied for a given vowel nucleus. The synthetic syllables (without bursts) are played for English-speaking listeners who are asked to identify the initial stop. In general, the results of these studies are fairly comparable, so I will relate the findings as a group.

The F2 transition for a /p/ or /b/ is rising into all vowels. The locus has been estimated as low as 720 Hz and as high as 1300 Hz. The F2 transition for a /t/ or /d/ is steady into front vowels and falling into back vowels. The locus is pretty much agreed to be around 1800 Hz. The F2 transition for /k/ or /g/ is falling into front vowels and steady or slightly falling into back vowels. There does not seem to be a single locus for the velar stops, rather there are at least 2 loci: a high locus around 3000 Hz for front vowels and a lower locus for back vowels.

The other approach to formant transitions, locus equations, seeks to determine an invariant relation between the actual beginning of a formant and the midpoint of a vowel for a given place of articulation. Lindblom (1963) was the first to point out that F2, measured at the onset of the vowel, is a linear function of the F2 measured at the vowel midpoint for a given consonant. The linear function is defined for a consonant across vowel types. Locus equations have subsequently been investigated primarily by Sussman et al. (1991, 1993, 1995) and Krull (1988, 1989, 1991).

Locus equations are derived by plotting F2 at the first glottal pulse and F2 at the vowel midpoint for a given place of articulation. The consonant remains constant while the following vowel changes for each type (e.g., [bi bɪ be bɛ bæ bɑ bɥ bo bʊ bu bʌ]). A regression line is then fit to the scatter plot of all F2 onset-F2 midpoint coordinates across vowel contexts. The slope of the regression provides an index of coarticulation. For example, a slope of 1 would indicate maximum coarticulation: the F2 at the glottal pulse and the F2 at the vowel midpoint would be perfectly positively correlated. The locus equations read:

$$Y = a + bX$$

Y is the predicted value of the F2 onset
a is the y-intercept
b is the slope
X is the value of the F2 vowel

From the many cross-linguistic studies with locus-equations, it can be determined that the labial stops have a steeper slope than the dental/alveolars. In other words, in the case of the labial, F2 is fairly near its eventual target frequency at the release of the stop. The dental/alveolar consonants have a flatter slope. The difference in slope has been considered to index the amount of coarticulation inherent in the place of articulation: labials have a steeper slope and are therefore more coarticulated than alveolars.

The velars present an interesting situation. Two locus equations are needed for them: one for the front vowel variant and one for the back vowel variant. The two types of velars are so different in place of articulation that one regression line put through the F2 onset-F2 midpoint coordinates would have a rather poor fit. Two regression lines handle the data much better. The velars before front vowels have a flat slope. In fact Sussman *et al.* (1991) found that the mean slope for the velars before front vowels was even flatter than that for the dentals. In other words, the vowel onset is fairly stable across the front vowels. The slope of the velars before back vowels has a steeper slope. This slope has been found to be even steeper than that for the the labials (Sussman *et al.* 1991). The Figures 2.1-2.4 provide example locus equations for [b], [d], [g]+front vowels, and [g]+ back vowels for single speakers. The figures are stylized versions of the actual locus equations presented in Sussman *et al.* (1991). Note that the [b] locus equation is steeper than the [d] locus equation and that the [g]+front vowel locus equation is quite flat while the [g]+back vowel locus equation is very steep. Also note the similarity between the locus equations for [g]+front vowel and [d] in that they both have a flat slope with a rather high y-intercept.

Example [b] Locus Equation

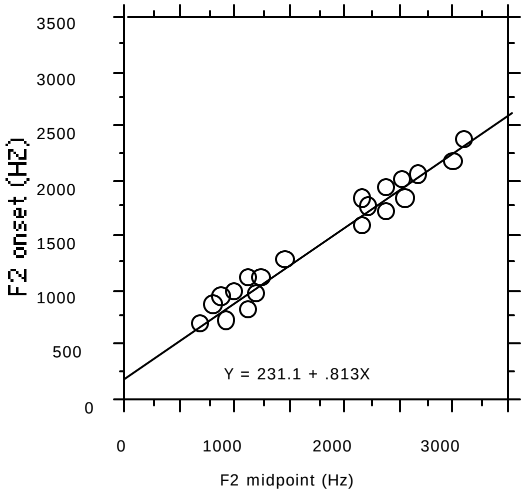

Figure 2.1

Example [d] Locus Equation

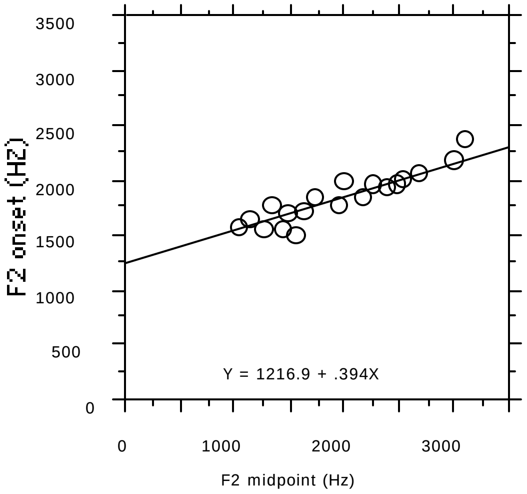

Figure 2.2

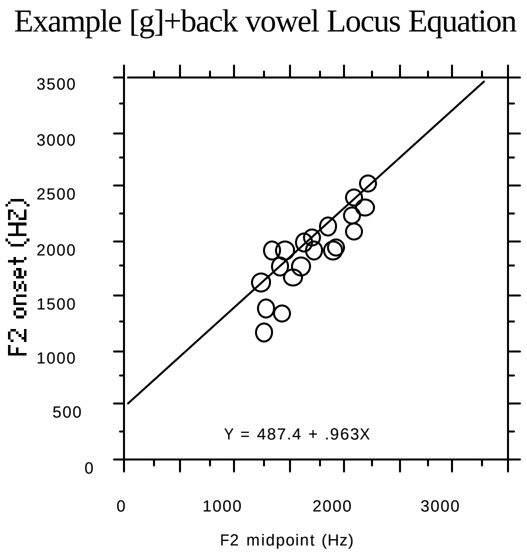
Example [g]+back vowel Locus Equation

Figure 2.3

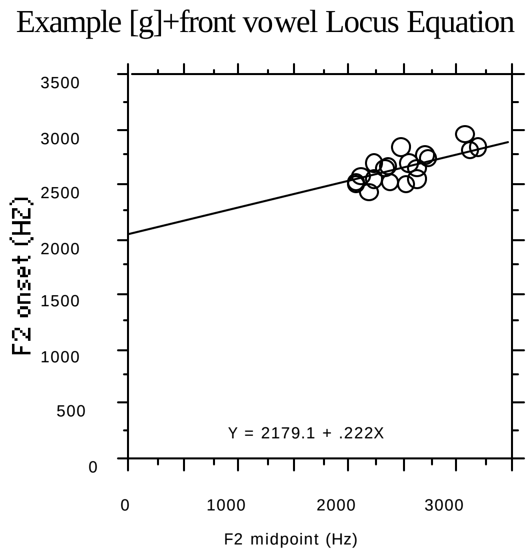
Example [g]+front vowel Locus Equation

Figure 2.4

Krull (1989), Duez (1989, 1990) and Poch-Olivé *et al.* (1989) have also investigated the slope of locus equations in spontaneous and elicited speech for Swedish, French, and Spanish and Catalan respectively. They reported that the locus equations have greater slope in spontaneous speech due to the fact that the locus and the nucleus are closer to each other, thus indicating more coarticulation and/or formant undershoot. Krull and Duez only look at the dental consonants [d n l] and the labial consonants [p m] in their analysis. They find that the spontaneous locus equations have a greater slope for each speaker in each place of articulation. It is also interesting to note that even as the slopes get steeper in the spontaneous speech, there is no overlap in the two places of articulation: the slopes for dental and labial remain distinct. Poch-Olivé *et al.* consider the voiceless stops [p t k] in Spanish and Catalan and find that overall, the slopes are greater in spontaneous speech. For the velars, however, the difference in slope is much smaller than for the other places of articulation. Based on this observation, they suggest that the velar stops behave differently.

2.4.1.3 Summary of the Acoustic Literature

In summary, we find that velars show more vowel context effects than other places of articulation. Evidence for this is found in articulatory and acoustic investigations. Velars also have a burst spectrum characterized by a single prominent peak which varies as a function of the following vowel. Labial and dental/alveolar burst spectra, on the other hand, are more diffuse and show less variation as a result of vowel context. Velars also have longer durations for burst and aspiration periods than other places of articulation and stops before high vowels have longer burst and aspiration intervals. The second formant transition shows more variation for velars than other places of articulation. There is not a single virtual starting point for velars in all vowel contexts as has been claimed for the alveolars. A locus equation analysis also finds it necessary to use two locus equations for velars as opposed to the single one needed for other places of articulation. In addition, the locus equations for velars before front vowels are quite similar to locus equations for dentals in that they both have a flat slope and a rather high y-intercept.

2.4.2. Perceptual Properties of Fronted Velars

Many perceptual studies have been conducted on burst information. Some of these studies have involved the manipulation of synthetic stimuli (Cooper *et al.* 1952, Liberman *et al.* 1952, Stevens and Blumstein 1978, Blumstein and Stevens 1980, and Hawkins and Stevens 1987), whereas others have used human stop bursts as stimuli (Halle *et al.* 1957, Malécot 1958, Winitz *et al.* 1972, Dorman *et al.* 1977, Tekieli and Cullinan 1979, Repp and Lin 1989, and Smits 1995). A common finding of the natural language studies is that a velar before a front vowel is often inaccurately perceived as a coronal (Winitiz *et al.* 1972, Repp and Lin 1989). This provides evidence that a velar to coronal sound change may have a perceptual basis. Evidently, the burst and aspiration information from a velar before a front vowel is perceptually similar to that from a coronal before front vowel. This correlates with the acoustic findings discussed earlier. The velar spectral peak before front vowels is high and thus more like an alveolar than the spectral peak of a velar before other vowels. In addition, the burst of a velar stop has been found to be more heavily weighted as a perceptual cue than the following formant transitions (Dorman *et al.* 1977, Smits 1995). Thus, it is possible that confusion of velar burst with a coronal burst would not be overridden by differing formant transitions. Another finding from studies using synthetic stimuli is that the absence of burst information was found to greatly impair the perception of place, especially in velars (Stevens and Blumstein 1978, 1980). So, formant transitions alone are not enough to accurately identify stop place of articulation for velars.

Perceptual studies using isolated bursts from human speech indicate that the information contained in the burst is enough for fairly accurate identification of place in stop consonants. Tekieli and Cullinan (1979) found that the first 10 ms. of an English word initial stop burst was all that was needed for a better than chance identification of place of articulation. Winitz *et al.* (1972) isolated the burst and aspiration (if present) of /p/, /t/, and /k/ release. The /p t k/ were taken from English words read in a sentence in which they were word initial or word final, either preceded or followed by /i/, /u/ or /a/. The bursts (and aspirations) were presented to listeners. Initial releases were identified better than final ones. The vowel was also identified reasonably well given the burst stimulus. The stimulus containing the burst and aspiration of /k/ before /i/ had a high error rate (81%); it was most often misheard as a /t/. From these findings, we can see that burst/aspiration information of a [k] before front vowel is highly confusable with a coronal place of articulation. If a listener were not attending to the formant transitions (as was the forced case here) of a velar before a front vowel, it could be misheard as a coronal.

Repp and Lin (1989) isolated the transient of English /b/, /d/, and /g/ before a variety of vowels from whispered speech. The talker kept the volume of air flow at a minimum to avoid any frication or aspiration. Listeners were able to identify both the consonant and the vowel from the isolated transient at a rate above chance, but not very accurately. Identification of [g] was good before back vowels, but heard as /d/ about 75% of the time before /i/ and /e/. Thus, a pure transient cue (i.e. devoid of frication or aspiration information) from a velar before high front vowel can be misheard as an alveolar.

Dorman et al. (1977) ran experiments which assessed the roles of release bursts and formant transitions as acoustic cues to place of articulation. They removed /b d g/ from onset position in syllables spoken in American English. They then cross-spliced them onto 9 English vowels which had been removed from syllables beginning with /b d g/. The results of perceptual experiments indicated that the velar bursts carried significant weight (as compared to the formant transitions) before the high front vowel /i/ as well as before the back vowels. They proposed that stop bursts carry more perceptual weight when their peak frequencies are close to the front cavity resonance found in the following vowel (F3 in the case of /i/ and F2 in the case of the back vowels). Smits (1995) also found that, in the case of velars, bursts were more important than formant transitions in consonant identification in Dutch. Smits' further findings indicate that the bursts of voiceless stops were more effective in cueing place of articulation than the bursts of voiced stops. Since the burst of velar before a high front vowel is weighted more heavily that the formant transitions, a velar burst misheard as a coronal might not be counteracted by (possibly) different formant transitions.

I will now review the studies which used synthetic stimuli. The findings indicate that high frequency bursts are heard as alveolars. Burst information is also found to be crucial in the identification of stop place of articulation.

Cooper *et al.* (1952) synthesized 12 bursts 15 ms. in duration ranging from 360-4320 Hz. These bursts were then matched with synthetic vowels composed of two steady state formants. They report that high frequency bursts (~2500Hz or above) were heard as /t/ for all vowels. Bursts were heard as /k/ when they were on a level with, or slightly above F2. Otherwise, the bursts were heard as /p/.

Stevens and Blumstein (1978, 1980) provide experimental results which indicate that burst information is necessary for accurate stop place identification. They also report that velars before high front vowels take longer to identify. Stevens and Blumstein (1978) combined various synthesized bursts with synthesized /i/, /u/, and /a/. The vowels had formant transitions systematically varying from /b/ to /d/ to /g/ and the bursts were correlated with the frequency of the varying formants. /b/ had the shortest burst (5 ms.) and /k/ had the longest burst (15 ms.). Listeners were instructed to identify the consonant . There was a 90% mean identification rate for stimuli with burst and transition information, 81% mean identification rate for transitions alone, and an 18% identification rate for bursts alone.

Blumstein and Stevens (1980) continue this line of investigation by presenting the first 10, 20, 29 or 46 ms. of synthetic CV's (beginning with either /b/, /d/, or /g/ and followed by /i/, /u/, or /a/) to subjects for consonant identification. The following vowels either had steady state formants or had vowel transitions considered to be optimal for /b/, /d/, and /g/. Bursts were present in some stimuli, but absent in some stimuli. The bursts were synthesized as follows. The burst for /b/ was 5 ms. long and equal to the frequency of F2 at vowel onset. The burst for /d/ was 10 ms. long and equal to the frequency of F4 before /a/ and F4 and F5 before /i/ and /u/. The burst for /g/ was 15 ms. long and equal to the frequency of F2 before /a/ and /u/ and F3 before /i/. For most stimuli, the first 10 to 20 ms. were sufficient to get reliable correct responses. This held true for stimuli with or without burst, and with formant movement or without formant movement. Identification of /gi/, however, was problematic. A /gi/ with burst (with or without formant transitions) was only reliably identified with the 46 ms. stimulus. A /gi/ without burst (with or without formant transitions) was poor for all lengths of stimuli.

To sum up, the perceptual experiments discussed here using stop bursts from human speech indicate that velars before front vowels are often confused with alveolars. The experiments using synthetic stimuli indicate that high frequency bursts (above 2500 Hz) are heard as coronals and that the velar plus front vowel combination is hard to identify correctly.

2.5. Acoustic Similarities of [k] and [tʃ]

If the sound change [k] > [tʃ] before a front vowel is perceptually motivated, then we would expect acoustic similarities between [k] before a front vowel and [tʃ]. What we find is that there are similarites between the burst and aspiration of a [k] before a front vowel and a [ʃ]. Here I review studies of acoustic cues to fricative place of articulation with particular attention to [ʃ]. As might be expected, the burst of [k] is spectrally similar to [ç] as well. In Chapter 5 I will outline why I think a fronted [k] more often becomes [tʃ] than a palatal affricate. Here it will suffice to note the acoustic similarities between fronted [k] and [ʃ] as well as [ç]. The literature on the perception of fricatives will also be discussed here. A generalization emerging from the studies is that [ʃ] (as well as [s]) are distinguished primarily by their noise portion and that the following vowel transitions are relatively unimportant. Thus, I hypothesize that if a listener somehow heard the noisy release of a [k] in a [ki] sequence as [ʃ], s/he could reinterpret that noise as a [ʃ] even if the transitions were not identical.

Before considering the spectral peaks of the fricatives studied let me briefly review the studies dealing with the spectral peaks of a fronted [k] burst. Halle, Hughes, and Radley (1957) found that the first 20 ms. of a [k] + front vowel has a peak from 2-4 kHz. Zue (1976) found that a velar before a front vowel had an average peak spectral frequency of 2720 Hz. Fischer-Jørgensen (1945) and Keating and Lahiri (1993) found that a fronted [k] has a spectral peak correlated with the following F3. In several articles discussed above Stevens and Blumstein found that a burst peak frequency around 3 kHz produced the best responses for the synthetic CV [gi] stimuli.

Many acoustic studies have been conducted on the spectral characteristics of fricative consonants. Here, I review and summarize the findings of these studies. Table 2.6 briefly summarizes the findings. The text that follows gives more detail.

Spectral Characteristics of Fricatives

Study	[s]	[ʃ]	[ç]	[x]
Hughes & Halle 1956 Behrens & Blumstein 1988	3.5-5kHz	2.5-3.5kHz		
Nartey 1982	1.5 & 6kHz	2 & 3kHz	1 & 3kHz	1.5 & 3.5kHz
Strevens 1960	3-4kHz	3-4kHz	3-4kHz	1.5kHz & 4-5.5kHz
Glave 1974			2.4-5kHz	.5-1.5kHz & 3.2-3.9kHz
Jassem 1965	above 4kHz	1.5-8kHz	2-8.5kHz	below 3kHz

Table 2.6

Studies of English fricatives (Hughes and Halle 1956 and Behrens and Blumstein 1988) have found that [s] shows a major frequency peak in the 3.5-5kHz range, while [ʃ] has a major frequency peak in the 2.5-3.5kHz range. Note that this is in the same frequency range for the spectral peak of a fronted [k].

Nartey (1982) conducted a cross-linguistic study of fricatives. He collected three tokens of each existing fricative from 5 native speakers of each of the following languages: Hopi, Navajo, Papago, Pima, Zuni, Amharic, Egyptian Arabic, Hebrew, Japanese, Korean, Polish and Yoruba. He then looked at spectral peaks in a 22 critical band anlaysis which models the auditory system. A rough average across the languages studied indicated the following spectral peaks (the measurements given here are from Nartey's translation of the 22 auditory critical bands into Hz): [s] 1.5kHz and 6kHz, [ʃ] 2kHz and 3kHz, [ç] 1kHz and 3kHz, [x] 1.5kHz and 3.5kHz. In this study, both [ʃ] and [ç] have a spectral peak at around 3kHz, similar to the fronted [k].

Several researchers have also investigated the range of spectral noise in fricatives. Though there is not a complete consensus, some have found that [ʃ] has a relatively short noise span in the frequency range. This can cautiously be considered a similarity with the compact spectrum of the velars. Strevens (1960) looked at temporally elongated fricatives produced by English speakers and found that [s ʃ ç] had short spectrum covering a range from 3-4kHz with a progressively lower limit from [s] to [ç]. [x] was found to have a medium spectrum covering a range from 4-5.5kHz with lower peak formant around 1.5kHz.

Glave (1974) studied the spectral range of the German fricatives [ç] and [x]. He found that [ç] had a range from 2.4kHz to 5kHz (except after the front rounded vowels where the range was from 1.9kHz to 3.8 kHz). He also noted that the noise for [x] consisted of a lower band from .5kHz to 1.5kHz and a higher band from about 3.2kHz to 3.9kHz.

Jassem (1965) collected fricative spectra for English, Polish and Swedish. He found that, in general, [s z] have most of their energy in a region above 4kHz; [ʃ ʒ] have most of their energy from 1.5kHz to 8kHz; [ç] have most of their energy between 2kHz and 8.5kHz; [x] has most all of its energy below 3kHz.

In sum, the studies discussed here that look at peak spectral frequency for fricatives all have slightly different results depending on the languages studied and the methodology used. We can, however, recover a few generalizations. The fricative [x] seems to have two distinct peaks, one much lower than the other. The fricatives [s], [ʃ] and [ç] all have peak frequencies that fall roughly in the 2-5 kHz range. This range is similar to the 2-4 kHz range found for the peak spectral frequency of the velar burst/aspiration before a front vowel.

The next studies to be reviewed all find that [ʃ] and [s] are mainly perceived based the noise portion and not the following transitions. This is in contrast to [f] and [θ], whose place is mainly encoded by the following transitions. Heinz and Stevens (1961) performed experiments in which they synthesized fricatives with an electric circuit containing one pole and one zero (which lay one octave below the pole) which was excited by white noise. The resonant frequency varied through several values from 2000 to 8000 Hz. These synthetic fricatives were played to listeners in isolation and with a following synthetic /a/ vowel. They found that "/f/ is distinguished from /θ/ apparently on the basis of the transition of the second formant in the adjacent vowel, whereas formant transitions do not have an appreciable effect on /s/ and /ʃ/ responses, particularly when the intensity of the fricative is high." (p.596)

When Heinz and Stevens played the synthetic fricatives to listeners in isolation, responses were very clear for /ʃ/, /s/ and /ç/ with strong cross-over points. The responses were less clear for /f/ and /θ/. Poles at 2000 and 2500 Hz were heard as /ʃ/. A 3500 Hz pole was heard as /ç/. A pole 5000 Hz was heard as /s/. A pole 6500 Hz was heard as either /s/, /f/ or /θ/. Finally, a pole at 8000 Hz was heard as either /f/ or /θ/. Here, in this perceptual experiment, only the isolated fricatives [ʃ] and [ç] fall into the 2-4kHz range defined for fronted velars through acoustic analysis.

Heinz and Stevens then combined the same fricative resonances used for the isolated fricatives, those with poles at 2500, 3500, 5000, 6500 and 8000 Hz, with a synthetic vowel /a/. The locus of F1 was 200 Hz for all the stimuli. The F2 locus was either 900, 1700, or 2400 Hz. The relative intensity of the fricative to the vowel varied from 0 db, -5 db, -15 db, and -25 db. /ʃ/ responses are always associated with the resonant frequencies in the vicinity of 2500 Hz. Strong /ʃ/ responses are obtained for all loci when the intensity of the fricative is 0 to -15 db in relation to the vowel. For the -25db tokens, the locus of F2 needed to be 1700 Hz or higher. /s/ responses are obtained when the resonant frequency of the stimulus is above 3000 Hz for all loci. There are more /s/ responses at the 5000 Hz level when the /s/ fricative noise is stronger in relation to the vowel. /f/ and /θ/ responses are obtained for the 6500 and 8000 Hz fricative resonances when they have very low (-25 db) intensity. /θ/ was heard when the locus was 1700 or 2400 Hz. /f/ was heard when the locus was 900 Hz.

Harris (1958) also found that "the important cues for the fricatives /s/ and /ʃ/ are given by the noise but that the differentiation of /f/ and /θ/ is accomplished primarily on the basis of cues contained in the vocalic part of the syllable. Similar results were obtained for the voiced counterparts of these sounds." (p. 1) Harris recorded English fricative+vowel syllables and then spliced them together for all combinations. When /s/ or /ʃ/ friction was paired with any vocalic portion, the resulting stimulus was judged as /s/ or /ʃ/ respectively. Apparently, the frication of /s/ and /ʃ/ provide the necessary and sufficient cues for their identification, and override whatever cues may be provided by the vocalic portions. Most of the listeners tended to judge a syllable with /f-θ/ friction as /f/ when it had an /f/ vocalic portion, and as /θ/ when it had any other vocalic portion. The results of the voiced fricatives test are similar, but not as clear.

Another experiment conducted by Whalen (1991) suggests that formant transitions might play a role in [s] and [ʃ] perception. Even more interesting for our purposes, however, is that he found perceptual confusion with [ʃ] and [k]. Whalen spliced the fricatives [s] and [ʃ] together in varying proportions. The fricative combinations were then spliced onto the front of a word, at the end of a word, and intervocalically. In one set of experiments, the listeners were forced to chose either /s/ or /ʃ/. Effects of duration and transitions were observed, in additon to the fricative spectra. In another experiment he had "expert transcribers" transcribe the hybrid fricatives before vowels. They were not forced to pick /s/ or /ʃ/, just to transcribe what they heard. They tended to hear two segments for the hybrid fricatives. Interestingly enough, when the /s/ was short (~25 ms.) they heard a /t/ for it, and when the /ʃ/ was short (~25-75 ms.), they heard a /k/ for it.

The studies done by Heinz and Stevens and Harris and Whalen all suggest that coronal fricatives can be identified solely by the noise portion, without information from the following vowel transitions. This suggests that a listener could hear the release of a fronted velar as a coronal fricative or even affricate (which would seem to be possible based on the similarities of their peak frequencies) even if the vowel transition were not identical.

Many studies have shown that, in general, voiceless fricatives are longer than voiced and that sibilant fricatives are longer than non-sibilant. Thus [ʃ] would be one of the longest fricatives just as a fronted [k] has the longest burst and aspiration of the stops. Baum and Blumstein (1987) and Crystal and House (1988) have found that English voiced fricatives are shorter than the voiceless fricatives. There is, however, overlap between the duration of voiced and voiceless fricatives. Manrique and Massone (1981) found that voiceless fricatives are longer than voiced in Spanish.

Sibilant fricatives are also longer than non-sibilant fricatives. Behrens and Blumstein (1988) found the following average duration for English fricatives: [s] 175ms, [ʃ] 175 ms., [f] 149 ms., and [θ] 134 ms.. Jongman (1989) found the following means for English fricatives: [s] 188ms., [ʃ] 166 ms., [z] 152 ms., [f] 149 ms., [ð] 119 ms., [v] 113 ms., and [θ] 107 ms. Jongman also found that the entire frication was not needed for correct perception. [ʃ,z] only needed 30 ms., [f,s,v] needed 50 ms., and [ð,θ] needed the entire duration.

In sum, the acoustic work conducted on fricatives indicates that the peak spectral frequency of [ʃ] and [ç] are similar to the peak spectral frequency of the burst of a fronted velar, i.e., the peaks fall in the 2-4 kHz range. The perceptual work on fricatives indicates that [ʃ] and [s] are mainly distinguished by their spectral properties and not the following formant transitions, unlike non-sibilant fricatives such as [f] and [θ] whose identification relies on the formant transitions. As discussed earlier in this chapter, voiceless velars are more likely to become affricates as a result of palatalization; the sound change [k] > [tʃ] is more common than the parallel sound change [g] > [dʒ]. This finding may have its roots in the fact that voiceless velar stops have longer VOT's than voiced velars and, hence, greater noise period following the release. The longer noise period is more likely to be reinterpreted as a fricative release such as [ʃ].

2.6. Summary and Discussion

In this chapter, several important aspects of velar palatalization conditioned by a front vowel or palatal glide are discussed. First of all, the most common outcome of the sound change is for the velar to become coronal. In the case of the [k], the most common outcome is [tʃ], and in the case of [g], in addition to [dʒ], a glide [j] or a coronal fricative is also common. A less common outcome of velar palatalization is the acquisition of secondary palatalization. This is usually in conjunction with the formation of a system-wide set of palatalized consonants. The focus of the experimental work in the following chapters is on the most common type of velar palatalization, namely velar to coronal palatalization. Other typological observations about velar palatalization are (1) the voiceless velars are more likely to undergo palatalization, (2) the result of the palatalization is usually a new segment in the language, and (3) the conditioning environment of a following front vowel is not usually lost. However, the conditioning environment of a following palatal glide is sometimes lost.

Traditional explanations for velar palatalization which rely on articulatory conditioning are also discussed. These explanations are found to be problematic in that each discrete articulatory step needed to explain the sound change cannot be adequately motivated. Due to the failure of articulatory accounts and spurred by the suggestion of Ohala (1989), I propose that velar palatalization is perceptually conditioned. I propose that velar palatalization in which a velar becomes a coronal preceding a front vowel or palatal glide is conditioned by a perceptual reanalysis of the front velar. By hypothesis, velars are more fronted in faster speech than in citation speech, therefore I further propose that velar palatalization arises from perceptual reanalysis of faster speech. This proposal makes several testable predictions: (1) velars before front vowels will be acoustically similar to coronals, especially the voiceless velars, (2) that similarity will be heightened in faster speech, and (3) fronted velars will be perceptually confusable with coronals. The acoustic and perceptual work discussed in this chapter provides preliminary support for these predictions: (1) fronted velars are found to have a peak spectral frequency of the burst similar to the peak spectral frequency of [ʃ] and [ç], (2) the formant transitions of velars before front vowels are also found to be more similar to dentals/alveolars than the formant transitions of velars before back vowels, and (3) velars before front vowels are found to be confusable with dentals/alveolars in perception studies.

The proposal that velar palatalization is conditioned by perceptual reanalysis of velars before front vowels in faster speech is tested more precisely in the experimental work discussed in Chapters 3 and 4. In Experiment 1, discussed in Chapter 3, I investigate the prediction that fronted velars and coronals are acoustically similar, especially in faster speech. I will focus on one of the most common outcomes of the sound change, namely the palatoalveolar affricate. I use the locus equation methodology to compare the F2 transitions of velars and palatoalveolars. I also investigate the spectral properties of the burst and look for similarities between fronted velars and palatoalveolars. In this study, two speech styles (citation and faster) are used to determine whether or not the faster speech tokens of the velars are more like palatoalveolars than the citation speech tokens. In Experiments 2 and 3, discussed in Chapter 4, I report the results of a perception experiment which investigates the confusability of velars and palatoalveolars in different vocalic environments.