The Phonetics of Labialized Velars in Ancient Greek

Helena Beeley

New College

University of Oxford

A thesis submitted for the degree of

Doctor of Philosophy

Faculty of Linguistics, Philology & Phonetics

Trinity 2016

Abstract

This thesis investigates conditioned sound changes of Proto-Indo-European labialized velars $*k^w$, $*g^w$, and $*g^{wh}$ to labials, coronals, and velars in Ancient Greek. Acoustic and perceptual experiments in British English and Western Zapotec provide typological evidence to inform conclusions regarding the phonetic inception of the sound changes.

Before back vowels and consonants PIE labialized velars became labials in Ancient Greek. Two competing hypotheses have been proposed for these sound changes: perceptual confusion, in which e.g. $*k^w$ developed to /p/ via misperception due to acoustic similarity between the two stops (Ohala, 1989, 1993); and incremental articulatory change, in which e.g. $*k^w$ gradually developed to /k͡p/ and then /p/ (Whatmough, 1937; Garrett & Johnson, 2013). The acoustic studies here found little acoustic or perceptual similarity between /kw/ and /p/ in any vocalic environment, indicating that the latter explanation is more plausible.

Before front vowels PIE labialized velars became coronals. Philologists have proposed developmental pathways involving processes of incremental palatalization. This thesis finds evidence for coarticulatory fronting of the velar articulation of labialized velars in front vowel environments, but the effect is language-specific and its auditory prominence depends upon the durational overlap of the primary and secondary articulations. Increased auditory prominence of coarticulatory fronting in mid vowel environments may explain divergent developments of $*g^{wh}$ in Greek to labials before front vowels and to coronals before front mid vowels.

Before the back round vowel /u/ the PIE labialized velars lost contrastive labialization in Greek. The acoustic and perceptual experiments support a perceptual reanalysis account, in which contrastive secondary labialization of the velar is reinterpreted as coarticulatory rounding adjacent to a rounded vowel.

Contents

1 Introduction

1.1 Statement of purpose

The purpose of this thesis is to identify the phonetic motivations for conditioned sound changes of the Proto-Indo-European labialized velars to labials, coronals, and velars in Ancient Greek.

Proto-Indo-European (PIE) is reconstructed with a series of labialized velar obstruents: voiceless unaspirated $*k^w$, voiced unaspirated $*g^w$, and voiced aspirated $*g^{wh}$, usually understood to be monophonemic stop + glide sequences e.g. [$\mathrm{k^w}$w], [$\mathrm{g^w}$w], [$\mathrm{g^{hw}}$w] (see e.g. Fortson 2010, p. 48–55). This series is attested in Mycenaean Greek (the earliest form of written Greek) but during the period between Mycenaean Greek and the first attested alphabetic Greek the labialized velars merged with other stops, in sound changes that were conditioned by the following vowel (Table 1.1).

These sound changes are well-known within the field of Proto-Indo-European philology, but as yet there has been no satisfactory explanation for their inception or path of development. The labial outcomes ($*k^w$o > /po/, $*k^w$a > /pa/

Table 1.1: A summary of the conditioned developments of PIE labialized velars in Greek, with selected examples. For an overview of the outcomes see e.g. Lejeune (1972, p. 43 ff.), Sihler (1995, p. 150 ff.).

PIE	Greek	Conditioning environment	Example
	p	_[a, o, C]	$*k^w$*oter-* > *póter-* 'which'
$*k^w$	t	_[e, i]	$*k^w$*e* > *te* 'and'
	k	_[u] *and* \[u]_	$*g^w$*ouk*^{w}*ol-* > *boukol-* 'cowherd'
	b	_[a, o, C, i]	$*g^w$*ī-* > *bio-* 'life'
$*g^w$	d	_[e]	**sm-g*^{w}*elb*^{h}*-* > *adelp*^{h}*-* 'sibling'
	g	_[u] *and* \[u]_	**h*$_2$*iu-g*^{w}*ih*$_3$*s* > *hugiḗs* 'healthy'
	p^h	_[a, o, C]	$*g^{wh}$*on-* > *p*^{h}*ón-* 'murder'
$*g^{wh}$	t^h	_[e, i]	$*g^{wh}$*erm-* > *t*^{h}*erm-* 'warm'
	k^h	_[u] *and* \[u]_	**l̥ng*^{wh}*-u-* > *elak*^{h}*ú-* 'light'

etc.) have parallels in multiple unrelated languages, suggesting a physiological motivation (i.e. in production or perception) independent of language-specific structural or phonological constraints; this has prompted theories that perceptual confusion of labialized velars and labials, due to their acoustic similarity, could have caused phonological misidentification which led to sound change (Ohala, 1993). An alternative explanation proposes gradual increase of the labial constriction, leading to intermediate articulations such as [k͡p] and [g͡b] in a chronological sequence $*k^w$ > /k͡p/ > /p/ etc. (Whatmough, 1937; Garrett & Johnson, 2013). The relative phonetic plausibility of these competing explanations has not yet been investigated. For the coronal outcomes ($*k^w$e > /te/, $*k^w$i > /ti/ etc.), some philologists have suggested incremental palatalization of $*k^w$ in front vowel environments, but there has been little research into how coarticulatory fronting of a velar closure in the context of front vowels may interact with the need for a posterior tongue position during the articulation of secondary labio-velarization. Loss of contrastive labialization in round vowel environments ($*k^w$u > /ku/ etc.) is typologically common and may be explained by 'hyper-correction' (as defined by Ohala 1993) in which secondary labialization of an obstruent is re-interpreted as the result of coarticulation, caused by a following round vowel. However, contrastively labialized stops are generally under-represented in perception experiments, leaving the mechanism somewhat unclear; similarity of articulation may also play a role. Thus, while these diverse and sometimes competing explanations are all intuitively plausible, and the sound changes have parallels both in other Indo-European languages and in unrelated language families, the mechanisms behind them are not yet well understood.

The problems posed by the sound change can be broken down into three main research questions: 1) What are the mechanisms behind conditioned sound changes of PIE $*k^w$, $*g^w$, and $*g^{wh}$ in Ancient Greek? 2) Are previous suggestions that perceptual confusion may have prompted the sound change of labialized velars to labials plausible? And 3), why are some of the sound changes described in Table 1.1 more typologically common than others? For example, sound changes of /kw/ to /k/ and /kw/ to /p/ are common cross-linguistically in different vowel environments, but changes such as /kw/ to /t/ before a front vowel are typologically very rare.

Questions 1) and 2) are approached with reference to previous proposals that a universally applicable perceptual confusion based on acoustic similarities between /kw/ and /p/ was the primary motivation for the sound change (Ohala, 1989; Garrett & Johnson, 2013). The validity of this hypothesis will be assessed by analysing the acoustic characteristics of labialized velars relative to the plain labial "outcome" stops using data from contemporary languages, and by provoking perceptual confusion in listeners to test the degree to which acoustic similarity (if indeed it does exist) corresponds to perceptual similarity between these sounds. An alternative account, proposed early in the history of Indo-European studies (Whatmough, 1937) and possibly supported by secondary evidence from modern spoken languages (Monzón & Seneff, 1984), postulates a progressive increase in the degree of constriction at the lips; this would convert labialization from secondary to primary articulation via successive intermediate stages of development including articulations such as [k͡p]. Research question 3) will examine typological evidence from a number of languages, with a focus on Ancient Greek as an example of a language showing typologically uncommon developments of labialized velars. Numerous “incremental” articulatory accounts have been suggested for sound changes of the PIE labialized velars to Greek coronals, and these will be analysed with reference to the acoustic and articulatory properties of relevant sounds in contemporary “proxy” languages; inscriptional and other historical evidence is taken into account.

Many previous approaches to these sound changes are based primarily on general descriptive phonetic considerations or on language-internal, structural accounts. In contrast, this thesis uses primary data from living languages and speaker/listeners, selected and analyzed specifically to address these historical linguistic problems. Four corpus-based and laboratory experiments support a detailed analysis of acoustic variation and perceptual biases in living speakers, which is then used to present an empirically supported account of each conditioned sound change.

1.2 Sound change

Recent approaches to sound change have focussed on natural diversity within speech communities and the role of individual variation, whether in perception, articulation, lexical choice, or syntax. The diversity of intra-speaker microvariations—and their presumed occasional progression to what might be termed macrovariations or sound changes—has been recognised in part thanks to the increased availability of precision apparatus to record and analyse speech sounds. The contribution that empirical research could make to the study of linguistic variation and change has become increasingly recognized as technology has become more accessible and more widely available. This represents in some ways a resurgence of pre-structuralist approaches to phonology, in that experimental phonology represents a “bottom up” approach to investigating linguistic structures. The starting point is audio and articulatory data from which theories about higher level structures may be inferred, in contrast to structuralist and generative approaches in which high level structures were considered primary and phonetic detail secondary or even incidental. Phonetic data has been increasingly exploited to provide explanations for perceived phonological structures in synchronic data (Lindblom, 1986, 1992; Kawasaki-Fukumori, 1992; Maddieson, 1996; Ohala & Ohala, 1992); for example, De Boer (2000) uses computer simulation to demonstrate how similar vowel systems arise through self-organization of computational agents which have been provided with realistic articulation and perception capabilities. The applicability of the experimental phonology approach extended from synchronic to historical phonology when the focus shifted to the role of articulatory and perceptual factors in contemporary variation (Ohala, 1971, 1974, 2005; Pagliuca & Mowrey, 1987; Chen & Wang, 1975; Kingston, 1992). For example, Beddor (2009) uses measurements of the duration of nasals and the temporal extent of their coarticulatory effects on preceding vowels to describe how coarticulation and perception interact to influence historical sound changes of a vowel followed by a nasal to a phonologically nasalized vowel. Similarly, Harrington (2012) uses an empirical analysis of German high back vowel variation to explain the actuation of British English high back vowel fronting. By acknowledging the importance of the speech signal in directing the course of linguistic change it is possible to compare typologically similar sound changes: similar changes in unrelated languages may be motivated by physiological characteristics of the human vocal tract, aerodynamic capabilities, or the neurological processes of perception. For example, Guion (1996) finds coarticulatory motivations for palatalization sound changes occurring independently in many unrelated languages, by identifying acoustic similarities between fronted velars and palatal stops which are the result of the physiology of the vocal tract rather than any language-specific phenomenon.

This thesis orients itself within the tradition of laboratory phonology and recognises the importance of signals as the primary source of linguistic data. Terms for higher level, abstract structures of language will be used for convenience where required, but are intended primarily for descriptive purposes. The mental storage method and retrieval process of these structures remains controversial, so the present project endeavours as far as possible to describe speech sounds, language-specific contrasts, and speaker-specific activities in terms of their phonetic realization and physiological bases. The sound changes in question are conditioned by a following vowel, indicating that the motivation for their development may lie in coarticulatory effects at the level of the speech signal. Some theorists argue that lexical items and frequent expressions are stored as exemplars, in which case the stored items are auditory and/or motor memories of dynamic speech resulting from connected, overlapping articulatory processes (Coleman, 1998; Välimaa-Blum, 2009) rather than as sequences of phonemes (although Pierrehumbert 2001, p. 148 argues that there are exemplar clouds for phonemes as well as for lexical items). The cognitive reality of the phoneme has been called into question by empirical studies such as Morais et al. (1979), in which adults who had never been taught to read alphabetically were far less able to decompose words into strings of constituent phonemes in the same way that alphabetically literate adults could, suggesting that the ability to segment speech into phonemes is learnt rather than innate. Read et al. (1986) observed similar results for alphabetic and non-alphabetic literate adults in China. Mann (1986) found that Japanese schoolchildren learning a syllabic writing system had less awareness of phonemes than American children of the same age, but both groups had an awareness of moras. Morais et al. (1987) found that non-readers could learn to segment words through experimental tasks, and suggested that the capacity for segmental phonological awareness can be accessed through conscious training and does not generally arise spontaneously. Port (2010) argues against the phoneme on the basis that human memories, as demonstrated in non-linguistic cognitive research, are richly detailed and episodic rather than purely category driven. Whilst memories may cluster or connect with one another to allow categorization, abstract generalizations are not necessary for comprehension; within this model, humans are capable of generalizing segmental units, and do so in order to perform specific tasks, but the absence of any consistent identifiable physical correlates for each context-independent segment imply that they are neither essential nor habitual for real-time processing (Port, 2010, p. 45). Instead, statistical generalizations are derived from past experiences; this framework has been effective in simulations of linguistic capacity (Grossberg & Myers, 2000).

Perhaps the strongest argument against the primacy of high-level phonological abstraction is that its invocation is rarely necessary (Browman & Goldstein, 1992; Blevins, 2004; Ohala, 2005; Silverman, 2006). Explanations offered in terms of phonological abstractions for specific phenomena can often be accounted for in parallel by exemplar models, or by analyses of aerodynamic constraints and close examination of the capabilities of the speech organs. Indeed, isolation of motor processes involving speech from those involving other human activities seems unnecessary; swapping of segments to produce linguistic errors due to fast speech need not be the result of different psychological processes than the production of errors in co-ordination of other complex motor movements, such as fingering mistakes when playing the piano, and as such need not be described in terms of the interaction of artificially segmented abstract units. Within this framework, the phonological patterns of a language emerge as a consequence of the interaction of physiological and cognitive pressures (Silverman, 2006, p. 186) rather than vice versa. Exemplar theories also provide space for morphological, syntactic, and semantic influences on sound change. Phonological changes are frequently influenced by higher-level structures, for example in the generalization of particular features within grammatical paradigms in which morphology and phonology influence one another, or 'folk' etymologies in which speakers semantically reanalyze compounds.

1.3 Variation and change

Although contemporary arguments in favour of separating processes of synchronic variation and diachronic change have some currency (Kavitskaya, 2002), the variationist perspective is embodied by the emphatic statement 'change is variation' (Labov, 1982). This theoretical approach blurs the distinction between synchronic and diachronic states of language; speakers have the capacity to switch between co-existing older and innovated forms, with change eventually occuring when the usage frequency of one variant relative to the other is sufficiently high to prevent its being heard or learnt by subsequent learners. Labov approached variation from a generative framework with the conception of the variable rule, in which variation selection in usage is governed by systematic rules (Labov, 1969) and outlined ways in which sociolinguistic methodologies could contribute to generative theory (Labov, 1972). Later research in sociolinguistics exposed the functionality of variation as an essential social and communicative tool (Eckert, 1988; Moore & Podesva, 2009), exposing ways in which language users consciously and unconsciously selected varying forms of expression. Hay & Drager (2010) observe subconscious variation within the speech of individual speakers based on contextual priming during interviewing; conscious choice of variants as an expression of identity or group belonging is explored in e.g. Johnstone (1999) and Bell (2002).

It has been argued that, since not all variation becomes embedded as change, a distinction should be drawn between variation which is available to speakers and variation which becomes embedded as historical change (Ohala, 1990a; Stuart-Smith, 2004). This thesis deals specifically with the inception of variations which may not necessarily become embedded as change but which have the potential to. Hale et al. (2007) make a distinction between variation and 'microvariation': variation involves a difference in feature representation of a sound, whereas microvariation is defined as non-phonological variation introduced by physical processes, but nonetheless systematically conditioned by context.[1]. Instances in which featurally distinct but non-contrastive sounds alternate contextually or stylistically constitute variation, and in terms of sound change represent a 'later' stage of development.

[1]The term is also used in sociolinguistics to refer to small shifts in the usage rate of a linguistic variable across a period of discourse (Becker, 2009, p. 634)

Microvariations, on the other hand, represent the lowest level of possibilities for phonetically motivated sound change.

This research suggests a close relationship between 'microvariation', synchronic variation in language use (whether referring to variation between different generations of the same speech community, or between different geographically or socially positioned subsets of a speech community, or within an individual speaker themselves) and the diachronic embedding of particular variations within a language to the extent that other variations are no longer acceptable. Therefore no absolute delineation between variation and change will be posited in this thesis; rather, 'sound change' will refer to historical cases in which a variation became widespread in the past, making it difficult or impossible to describe the earlier state of variation. Identifying which of a range of possible variants might lead to phonological change in a specific linguistic situation, which variants might not, and why, represents a later stage of development in the history of any individual sound change; the first step is identifying phonetic variation which has the potential to lead to phonological change.

1.3.1 Coarticulation and 'correction' effects

The analysis of coarticulatory gestures offered a fruitful source of sound change actuation accounts, possibly because it is as an empirically observable cause of variants that often demonstrates similar effects cross-linguistically. All speech sounds overlap with those adjacent to them, and are often affected even by sounds occurring much sooner or later than them in the utterance; experimental data confirms that language users are capable of recognising and accounting for the effects of coarticulation in order to understand the speech of others (Mann & Repp, 1980; Shockey, 1977; Harrington et al., 2008). In Mann & Repp (1980) subjects perceived the same artificial stimuli, namely fricatives along a continuum from [s] to [ʃ], differently depending on which vowel followed the fricative - some stimuli that were identified as [ʃ] before an [ɑ] vowel were identified as [s] before an [u] vowel, demonstrating how participants compensated for the coarticulatory effects they expected to hear in the vicinity of different vowels. Elsewhere, listeners were able to recover [ð] despite its almost complete loss following [n], and in fact [ð] perception was induced by artificially elongating [n] in the phrase "warn a guy", meaning listeners reported having heard "warn the guy" (Shockey, 1977; Mann & Repp, 1980). Mitterer (2006) took the investigation a stage further by identifying multiple mechanisms for listener coarticulation compensation; listeners compensated for anticipatory lip-rounding in audio stimuli consisting of fricative-vowel sequences when also exposed to a visual cue to lip-rounding. Mitterer notes (p. 1239) that similar experiments performed with liquid-stop clusters found that visual cues did not induce compensatory effects, concluding that multiple interacting mechanisms allow listeners to resolve contextual effects.

The term "hypocorrection" was coined by Ohala (1989) to describe scenarios in which listeners, accustomed to accounting for natural coarticulation effects, fail to do so accurately. Instead, the articulatory effects of one sound are reinterpreted as articulatory characteristics of an adjacent sound. For example, French *bon*, produced as [bõn], might be heard and reanalysed by the listener as if it were [bõ]. Conversely, "hypercorrection" is defined as the overzealous application of the same corrective processes; articulatory movements are interpreted by the listener as evidence of the presence of a sound which the speaker did not consciously intend to articulate. For example, Slavic synchronic backing of [a] near palatalised consonants may seem counterintuitive as a sound change; fronting would be expected near a palatal, but a hypercorrection account proposes that listeners knew this and reconstructed a back vowel, assuming that the fronting was a coarticulatory effect of the preceding palatal (Ohala, 1989; Darden, 1970). The theory provides an explanation for many instances of dissimilation, and implies that articulations that are extended over long time intervals will be more liable to dissimilation; a cross-linguistic survey of synchronic dissimilation patterns suggests that this might be the case (Suzuki, 1998). Laterals, for example, have exceptionally long F_2 and F_3 transitions and are the most common articulatory category subject to dissimilation (Alderete & Frisch, 2007; Suzuki, 1998; Bye, 2011). The applicability of this theory to long-distance dissimilations has been called into question (Alderete & Frisch, 2007) but it is clear from phonetic evidence that coarticulatory effects can extend well beyond syllable and even word boundaries (Daniloff & Moll, 1968; Magen, 1997; West, 1999; Heid & Hawkins, 2000) and, if there were a need to propose alternative mechanisms for distance dissimilation based on phonological constraints, there is no reason why such mechanisms could not coexist with hypo- and hypercorrection. Beckman et al. (1992) explored the link between coarticulation and sound change by analysing the interaction of coarticulation and prosody, finding that compensatory reinterpretation of fast-speech processes can contribute to sound change.

The theoretical role of coarticulation-induced effects in sound change, whilst providing possible scenarios in which people may make individual processing errors by misinterpreting contextual cues, nonetheless fails to explain how these errors, manifested in one-to-one speaker-listener interactions, can result in community-wide sound changes. Language users have community norms, past experience, societal expectations, and in many contemporary communities writing conventions to refer to when processing speech input. Some approaches stress the primacy of the speaker, as opposed to the listener, in motivating sound change. Pagliuca & Mowrey (1987) argue that all sound changes in fact demonstrate reduction, or the removal of sounds or articulatory processes for the purposes of neuromuscular efficiency; according to this approach, there is no such thing as fortition or the addition of new linguistic elements. Browman & Goldstein (1990, 1992) describe phonology in terms of articulatory gestures, proposing that variation in production is the result of increase in overlap and/or decrease of magnitude of gestures. Phenomena such as the development of stops into affricates (as in some early Germanic consonant shifts such as **p > pf*) which might be described in feature-based terms as stops gaining a feature and moving from simple to complex articulation—and which in auditory terms results in an increase in acoustic energy—are described as the result of substantive reduction of stop closures. In contrast, coarticulation is the result of temporal reduction, with speakers overlapping gestures and undershooting targets in order to reduce the expenditure of neuromuscular energy. This obviates the need for acoustic explanations of sound changes; Pagliuca & Mowrey argue that a model based on individual misperceptions would result in a kind of wild oscillation between synchronically variable phonetic realisations and lightning-quick sound changes. However, discounting the role of listening leaves no room for explaining how this articulatory variation is processed and why language users elevate articulatory changes above simple efficiency and towards permanent phonological change; the role of perception and cognition may be minimised but cannot be outright ignored.

1.4 Perception and perceptual biases

Lindblom (1990) presents a theory of phonetic variation in which the tendency of speakers to economize muscular effort competes against the need of the listener to extract meaning from the signal. Ohala (1971, 1993) and Ohala & Lorentz (1977) investigate the role of speaker misperception in sound change, suggesting that a speaker's need for gestural economy may overwhelm a listener's ability to parse an utterance; however, the explanation of how individual misperceptions may lead to sound changes across languages and communities has never been satisfactory. Bybee (2010, 2012) proposes a significant modification to the "misperception" theory of sound change that helps to reconcile the auditory with the articulatory; her suggestions go some way towards eliminating these problems. Under Bybee's model, small changes in articulation prompted by language-specific prosodic patterns and coarticulatory gestures result in a tendency to reanalyse phonemes in a particular way; this differs from Ohala's theory in that the articulatory and perceptual changes are concurrent. Perhaps the most important aspect of this model is that all speakers in a given speech community are subject to the same tendencies; this specifically addresses earlier criticism of the apparent difficulty in leaping from individual, innovative "misperception" to regular and community-wide sound change. Ohala's theory gives an overly prominent role to language learners, a segment of the population whose influence in determining the spread of linguistic variants is not as certain as might intuitively seem to be the case (Bybee, 2010; Croft, 2000); on the other hand, the exemplar model accredits the potential to incept change to all speakers. A suitable illustration for this might be the development from [ɫ] to [w] in Romance languages (Recasens, 2012). The two sounds share similar acoustic characteristics — most notably a long, slowly-changing interval of voicing, with relatively low F_2 frequency (Ohala, 1974) — and an articulatory motivation prompted by undershooting the alveolar target is also eminently plausible. A sound change of this kind fits well into Bybee's suggested model in that acoustic similarity and articulatory tendencies could have worked in tandem to produce the given sound change.

Variation in the perceptual grammars of speakers has been demonstrated in the laboratory (Janson, 1986; Harrington et al., 2008; Beddor, 2012). In Janson (1986) an experiment on vowel perception in Stockholmers revealed variation in the amount of perceived backness of Swedish [aː], a disparity linked to the age of the participants. Participants were exposed to artificial stimuli replicating vowel sounds along a continuum from [oː] to [aː] (the frequencies of f_0, F_3, and F_4 were kept constant, with F_1 and F_2 frequencies decreasing in increments from the means of Stockholm production of [oː] to those of [aː]) and were then asked to identify which vowel they heard for each token. In terms of formant frequencies, the boundary at which [oː] became perceived as [aː] was lower for younger participants than for their older counterparts, indicating that younger speakers' perceptual space for [aː] is more retracted than older speakers'. Although Janson did not investigate whether or not this perceptual change was reflected in production, the study provides evidence not only that gradual sound change exists and that it is possible to observe it in progress, but also that the role of perception is of equal importance to that of production in the study of sound variation and change. Further work on perception in sound change has focussed on frequency effects and the concept of the exemplar cloud, whose centre of gravity shifts according to the token frequency of certain perceptions and so affects production (Harrington et al., 2008; Pierrehumbert, 2001). It has been suggested that speakers can exploit frequency effects presumed inherent to listeners' cognition processes to increase efficiency within individual conversations (Fowler & Housum, 1987; Hawkins & Warren, 1994). Of course, these tokens must be produced somewhere in order for them to be perceived and stored in the exemplar cloud; perhaps the important aspect is that the variant production is not stored in the exemplar cloud of the speaker before it is uttered, but once it is uttered it is part of the speaker's and the listener's inventory and stored for future reference.

Perceptual similarity as a result of acoustic similarity between the "original" sound and the result of the change has been an important focus of attention in the study of perception-based sound change (Guion, 1998; Blevins, 2004; Ohala, 1974; Mann & Repp, 1980; Pagliuca & Mowrey, 1987). Some typologically common sound changes, for example that from [θ] to [f], are difficult to explain in terms of articulatory undershoot or the confusion of complex motor processes; in this example articulation changes from an interdental fricative to a labio-dental fricative, a development which is not easily explained by articulatory undershoot or coarticulatory gestures. However, Miller & Nicely (1955) revealed a perceptual bias in confusion between English [f] and [θ]: confusion was weighted in favour of [θ] > [f] confusion (i.e. participants were more likely to select [f] as a response to a [θ] stimulus than vice versa). This is strongly reminiscent of cross-linguistic patterns in sound change, in which a historical change [θ] > [f] is typologically more likely to occur than a change [f] > [θ] (Johnson et al., 2011; Blevins, 2004). Observations such as this have led to research in which perceptual confusions are purposefully provoked with the intention of applying the results to phenomena of historical sound change. Although the laboratory is not a natural environment for sound change to take place, and provoking misperception by mechanically obscuring the speech signal hardly replicates normal speaker-listener experiences, this approach has proven effective. Guion (1996, 1998) compared the production and perception of velars and palatoalveolars in different vocalic environments and at different rates of speaking, successfully demonstrating that misperception could lead to phonological reanalysis of coarticulatorily fronted velars as palatoalveolars. Johnson (2003) links acoustic and perceptual similarity to sound change via the concept of perceptual "licensing", i.e. the perceptual similarity between two sounds makes listeners less likely to distinguish between them. Perceptual confusion accounts have been suggested for many other typologically common sound changes, including the development of labialized velar [k^w] to [p] (Ohala, 1989; Garrett & Johnson, 2013). A distinction is sometimes made in perceptual confusion accounts that perceptual errors do not lead in a direct way to sound change, but that perceptual confusion of a specific sound is subject to bias in a particular direction; in the model proposed by Bybee and described previously in this section all or some of the population may be subject to a perceptual bias and therefore likely to internalize biased representations. Wilson (2006) demonstrated that subjects generalized new patterns of palatalization more quickly in contexts where a demonstrable perceptual confusion bias was present than in contexts where such a bias was not observed. The combination of demonstrable perceptual bias, and the focus on the speech community as a whole as described by Bybee, helps to resolve some of the problems posed by Ohala's original model of perceptual misparsing (Hale, 2012), which seems at odds with the observable Neogrammarian hypothesis of sound change regularity (or near regularity); placing agency for the inception of sound changes into the hands of the speech community, rather than individuals, helps to explain how variants become capable of gaining social traction, whilst also going some way to explain the lack of "oscillation" predicted by Pagliuca & Mowrey (1987), on the basis that the speech community has the collective power to reinforce norms as well as to change them.

1.4.1 Labialized velars

The main focus of this thesis is on the sound changes involving labialized velars described in §1.1. The Greek developments represent both phonetic and phonological change; they resulted in neutralizations of contrast between labialized velars and labials, coronals, and velars in different vocalic environments. However, the thesis takes the view that the motivations behind the changes were primarily phonetic, rather than structural or analogical in nature. This is based partly on the fact that the divergent outcomes are conditioned by the phonetic environments, and partly on the observation of typologically parallel changes in unrelated languages; similar phonetic changes are observed to have taken place within very different phonological systems (examples will be discussed in Chapter 5). For these sound changes the problem of how phonetic variants are phonologized and subsequently extended out of the conditioning environment in which they arose need not be approached directly.

An acoustic/perceptual explanation for the development of labialized velar stops to labial stops has been proposed, based on evidence from a perceptual confusion study (Winitz et al., 1972). Ohala (1989, 1993) drew attention to the possibility that, given the comparatively high rate of perceptual confusion between English /k/ and /p/ before back rounded vowels, an acoustic/perceptually motivated sound change may have produced the Greek outcomes. The experimenters recorded instances of /p, t, k/ in running speech and in combination with each of three different vowels, /i, a, u/. Participants then heard a) the bursts of /p, t, k/ isolated from their context, and b) the isolated bursts with 100 ms of the preceding or following vowel. When presented with /k/ input followed by 100 ms of the vowel /u/, participants incorrectly identified /k/ as /p/ in 23% of trials. Labialized velars had already undergone delabialization in positions adjacent to $*u$ in an earlier stage of the language (see extended discussion in §2.6) but a comparison can be drawn with developments of labialized velars adjacent to the other Greek rounded vowel, in which $*k^w$ changed to /p/ adjacent to /o/. The vowels /o/ and /u/ differ in height but share relatively low F_1 and F_2 frequencies and lip rounding, suggesting it would be reasonable to propose a similar perceptual confusion pattern for the Greek sounds (see Chapter 2 for discussion of Greek vowel qualities). When hearing the stimulus consonants followed by 100 ms of each vowel, respondents were found to be more likely to mistake /k/ for /t/ ($p = 0.47$) than to accurately identify it as /k/ ($p = 0.38$) in the position immediately preceding /i/. This has been taken as evidence to indicate that there is likely to be a corresponding perceptual similarity between /p/ and /kʷ/ in the context immediately preceding a back rounded vowel, given the acoustic similarities between /k/ and /kʷ/ in this position, and given the evidence from Greek and other languages for developments of $*k^w$ to a labial before a back vowel (Ohala, 1989).

An alternative explanation involves incremental articulatory changes from labialized velars to labials. Garrett & Johnson (2013, p. 39) support a perceptual confusion account for sound changes of stops or fricatives with secondary labialization to labial stops or fricatives, such as /kʷ/ to /p/ and /xʷ/ to /f/ (as in Old English to Buchan Scots); however, they also describe an alternative articulatory explanation in which a /w/ off-glide becomes a stop or fricative, leading to an intermediate stage in which the articulation is [k͜p] or [x͜ɸ], after which the velar articulation is lost. The same articulatory process is described by Ponelis (1974, p. 28) as "narrowing", i.e. the narrowing of the constriction during the secondary labial articulation. This thesis assesses both of these possible explanations using acoustic and perceptual data.

1.5 Use of proxy languages

Typological inferences have always been used to some degree in historical linguistics; even the comparative method relies on judgements regarding the likelihood of specific sound changes (Kümmel, 2015). The uniformitarian hypothesis suggests that variation unobservable in ancient languages can be considered comparable to variation observed in currently spoken languages. Thus the range of possible articulations of a specific phoneme in a living language is likely to be similar to the range of possible articulations of the equivalent or similar phoneme in an ancient language. Several approaches have successfully used data gathered in laboratories to provide plausible explanations for sound changes in languages for which speech data is no longer available. For example, Stuart-Smith (2004) compares evidence from Latin authors and inscriptions to contemporary phonetic studies to explain how variation in production of aspirated stops may become phonologized and lead to sound changes of aspirated stops to fricatives. Kavitskaya (2002) uses evidence from the study of intrinsic vowel durations to propose a route by which lengthened vowels may have become phonologized in compensatory lengthening processes in Ancient Greek and other languages. Sen (2012) uses phonological syllable duration rules in Turkish and Finnish, derived from phonetic vowel reduction, to propose phonetic and phonological mechanisms by which Latin vowels could have been neutralized in open syllables. Guion (1996, 1998) uses data from native American English speakers to describe how cross-linguistically common processes of historical velar fronting can be explained as the result of perceptual shift. Mielke (2012) acknowledges that some phonetic details are unavailable when using proxy languages, but uses English speakers to describe ways in which phonetics may affect phonology cross-linguistically and makes the point that studying any language using recorded speech data necessitates abstraction away from the synchronic or diachronic reality of that language (p. 146, n. 1). These studies constitute surveys of the articulatory and acoustic characteristics of particular phonemically contrasting sounds in contemporary languages, which are then analysed with reference to extant written data from those ancient languages which are believed to have contained similar contrasts — realised in a phonetically similar manner — within their inventories. This enables "balance of probability" accounts of historical sound changes to be proposed. Typologically similar phenomena are catalogued to provide support for the proposed account on the basis that speech organs are universal, and consequently similar effects in unrelated and out-of-contact languages are likely to be the result of physiological and perceptual similarities. For example, Stuart-Smith (2004) compares intervocalic fricativization of voiced aspirates in Spanish to the fricativization of the inherited Indo-European voiced aspirates in Latin; she argues that the extremely short closure duration of breathy voiced stops, demonstrated by a review of experimental data, increases the likelihood of articulatory undershoot. This approach depends upon two premises: that sound change can be prompted by phenomena within the realm of phonetics, and that the uniformitarian hypothesis is true for human language. In this context the uniformitarian hypothesis proposes that variation found in speech production in the present is equivalent in nature and extent to variation found in languages everywhere, including those which are no longer spoken; likewise, phonetic constraints resulting from physiological characteristics of productive and perceptual speech processes apply equally in contemporary and historical languages (Labov, 1982; Croft, 2000; Stuart-Smith, 2004). Given that the reconstruction of labialized velars in Greek is generally considered secure (see Chapter 2), analysis of labialized velar acoustics and behaviour in currently spoken languages will allow the best estimate of the nature of the variation likely to have been produced by speakers of ancient Greek.

In short, there is strong evidence to suggest that phonetically based accounts of sound changes can provide successful explanations of sound change initiation and development, particularly in instances in which a physiological or auditory solution is made plausible by repeated observations of similar sound changes in unrelated or temporally distinct languages. This is not to say that traditionally phonological or generative accounts are not relevant in many instances; at some level listeners must interpret and categorise phonetic information to derive meaning, and many historical linguistic developments (such as push or pull chain sound changes) require reference to the phonological awareness of the speaker. However, there are sufficient examples of sound changes for which phonetic accounts offer realistic explanations to make it seem worthwhile investigating phonetic motivations for the Ancient Greek sound changes. This thesis will address the question of how Ancient Greek sound changes of labialized velars to labials, coronals, and non-labialized velars took place, taking into account typologically similar acoustic evidence from British English and Western Zapotec and perceptual evidence from British English speaker-listeners. The following section concludes the introduction by outlining in more detail how this thesis will pursue this investigation.

1.6 Structure of the thesis

Chapter 2: Labialized Velars in early Greek outlines the developments of labialized velars in Greek, beginning with an analysis of the reasons and justification for reconstructing labialized velars in Proto-Indo-European using the comparative method. Specific problems and controversies relating to labialized velars in Ancient Greek are outlined, with reference to Mycenaean Greek, Greek from the alphabetic period, and comparative evidence from other Indo-European languages. Greek labialized velar sound changes were conditioned by the quality of the following vowel; a discussion of vowel quality in Ancient Greek contextualizes the obstruent developments.

Chapter 3: Acoustic Characteristics of Labialized Velars provides a summary of the acoustic characteristics of labialized velars in some modern languages. A review of literature relating to the articulation and acoustics of labialized velars is presented. Results are presented from the acoustic analysis of labialized velars in different vowel environments from the Audio British National Corpus (Experiment 1), from British English material recorded for this study (Experiment 2), and from the Zapotec and Chatino Survey Archive (Experiment 3). A cross-linguistic comparison identifies parameters of acoustic variation of labialized velar stops in front, non-front, and rounded vowel environments.

Chapter 4: Perception of Labialized Velars describes a perceptual confusion experiment (Experiment 4) in which participants heard British English stimulus words obscured by background noise in order to provoke misperception. Subjects were not more likely to mistake /k^w/ for /p/ at a rate higher than for other stops, but they were very likely to mistake /k^w/ for /k/ in back round vowel environments. There was a significant correlation between the acoustic and perceptual distances between tokens.

Chapter 5: Phonetic explanations for the Greek sound changes presents a discussion of all three conditioned sound changes, taking into account empirical evidence from Chapters 3 and 4. Typological evidence from other languages is presented to support a phonetically motivated account of the sound changes. Results of the perceptual confusion experiment described in Chapter 4 are interpreted and indicate that a misperception process as outlined in Ohala (1989) is unlikely, despite acoustic similarities between /k^w/ and /p/ discovered in Chapter 3. Evidence from the acoustic study described in Chapter 3 is further analyzed to demonstrate that actual synchronic variation variation in labialized velar fronting before front vowels suggests potential for historical development. Evidence from Chapters 3 and 4 is taken into account to propose a perceptual reanalysis account of the Ancient Greek loss of contrastive labialization in back round vowel environments.

Conclusions are summarized in Chapter 6.

2

Labialized velars in early Greek

This chapter addresses the question of how we know what the 'starting points' of the sound changes were, and provides a description of the labialized velars as they were in Proto-Greek. Comparative evidence from Ancient Greek and from other IE languages is presented to identify details of the sounds and the system of contrasts that existed in the parent language, and to describe the conditioning environments for the sound changes. The philological background for the reconstruction of the labialized velars is outlined in this chapter, explaining why they are reconstructed and investigating the extent to which they can be described in articulatory and acoustic terms. An overview of the Greek reflexes of the labialized velars was provided in Table 1.1, although the picture is more complex due to differences between the dialects; the reflexes in Table 1.1 are those found in the Attic dialect (see e.g. Lejeune 1972, §30–31; Sihler 1995), and a detailed presentation of the developments in other dialects of Greek is provided in §2.5.

2.1 Pronunciation of dead languages

There are many challenges faced by historical linguists in tracing past sound changes and reconstructing ancestor phonemes; perhaps the first problem to approach in determining how completed sound changes were initiated is that of knowing how the 'original' sounds were actually produced. In the case of those sound changes which are ongoing, linguistically related speech communities which have not innovated in the same direction may provide usable information on the acoustic characteristics of the sound which underwent the change. Likewise, older and younger speakers may demonstrate differing phonetic realisations of phonemes, allowing for side-by-side comparison of the original and innovative sounds (Janson, 1986). In the case of languages in which no speakers retain any characteristics of the sound in question, such data is of course unavailable, and in some instances the sound change studied took place sufficiently far in the past that even sound change 'metadata' such as the quality of adjacent vowels, or of phonemically contrasting obstruents, cannot be extracted from contemporary speakers. The developments of labialized velars to coronals and labials occurred in at least some dialects of Ancient Greek before the beginning of the alphabetic period (Lejeune, 1972) in the eighth century B.C.E. (loss of contrastive labialization in back round vowel environments may have occurred much earlier, before the divergence of Greek from other Indo-European languages, §2.6) and so are of course of this latter type.

Naturally, secondary sources of the kind generally used to reconstruct the sounds of ancient languages do not allow us to establish exact phonetic values corresponding to the written signs that survive, including coarticulatory gestures and phonetic variation within speech communities. However, it is possible to use various types of evidence to give a rough account of the phonemic oppositions and structures within certain timeframes and geographic locations, and to assign some articulatory details to those phonemes we can identify. In this section types of evidence will be presented to describe the reconstruction of labialized velars in Indo-European and early Greek and to describe—to the extent possible—their phonetic and phonemic characteristics. It is also crucial to understand as fully as possible the characteristics of the outcome stops in later attested Greek, in order to draw conclusions about the processes of change, and to make comparisons with sound changes in other languages and with relevant phonetic evidence gathered from modern speakers.

Additional problems are posed by the fact that, whilst attested evidence is available to help in determining the qualities of the outcome sounds, the acoustic and articulatory qualities of the Indo-European labialized velars are determined entirely by comparative evidence from attested 'daughter' languages. Allowing for a certain margin of error, however, it is still possible to draw some inferences from this evidence; it is important to bear in mind that even when using data gathered from modern speakers in the most desirable conditions possible, there is still a level of abstraction between the researcher and the speech community, and all conclusions reached are constrained by the limitations of the evidence available. The following sections rely heavily on Allen's 1987 *Vox Graeca.*

Contemporary descriptions Whilst there are no extant descriptions of Ancient Greek phonology with the level of detail observed in, for example, Pāṇini's grammar of Sanskrit, there are some surviving descriptions of the sounds of Ancient Greek. For example, Dionysius of Halicarnassus, a Greek teacher of rhetoric writing in the latter half of the first century B.C.E, describes the coronal series as being produced with the tongue *katà toùs meteṓrous odóntas*, 'against the upper teeth' (*On Literary Composition* 14). Some ancient writers divided stops into groups, describing κ, π, τ as *psilós* 'bare, simple', whilst ϑ, φ, χ are *dasú* 'shaggy; hoarse' (Strato of Lampsacus, writing in the third century BCE, *De Audibilibus*, 804 b, 8-11; Dionysius of Halicarnassus, *De Compositione Verborum* 14; *Ars Grammatica*, attributed to Dionysius Thrax, 631.21). The Greek letters β, δ, γ are described as *metaxù toútōn* 'between these'. The description of these latter stops as something intermediate between the aspirated and unaspirated stops by both authors is unclear in its phonetic significance, and has resulted in some debate over their meaning. Whilst this evidence is useful, all of the authors cited here were writing several centuries after the first attestation of alphabetic Greek, by which point the pronunciation of many sounds had changed; for example, there is evidence that the aspirated stops had become fricativized for some speakers as early as the 1st century CE (Allen, 1987, p.21), with the result that one later discussion of the meaning of *psilós* and *dasú* attempts to apply them as terms descriptive of the difference between fricatives and plosives, as opposed to the difference between unaspirated and aspirated plosives (Allen, 1987, p. 23).

Evidence from within ancient languages It is possible to find evidence of ancient pronunciation within Greek, whether from synchronic alternations in Ancient Greek or by investigating later developments in Koine and Modern Greek. For example, Ancient Greek alternations such as λέγω, *légō*, 1.sg act. ind. 'say', and λελέκται, *leléktai*, 3.sg perf. mid. 'say', suggest that γ and κ differ in voicing, with κ representative of a voiceless velar sound resulting from assimilation to the following voiceless τ (Allen, 1987, pp. 27-30). Similarly, the patterning of ϑ, φ, χ with κ, π, τ, rather than with fricatives, in the application of suprasegmental rules in Attic verse suggests that they were plosives rather than fricatives (Allen, 1987, p. 19). Spelling differences between the dialects help to shed light on the order and type of sound changes undergone in each dialect; for example, the Boeotians did not use the υ grapheme to transcribe their short back round vowel when they adopted the Attic alphabet in the 4th century BCE, instead using a digraph ου, suggesting that the Attic pronunciation was by this point sufficiently altered that it no longer resembled the back round vowel of the Boeotian dialect; later, the Boeotians did use υ to represent Attic οι, suggesting an anterior articulation was associated with this grapheme (Allen, 1987). Despite numerous sound changes there is still some information to be gathered from Modern Greek; for example, Modern

Greek pronunciation of φ, ϑ, χ as voiceless labial, coronal, and velar fricatives respectively indicate—in the absence of evidence to the contrary—their historical places of articulation, if not their manner of articulation.

Spelling mistakes Scribal errors and confusion of letters often provide sources of evidence for changing pronunciations. For example, spelling confusion between Attic ο and ου from the 5th century BCE indicates merger of inherited diphthong /ou/ with the long round vowel /ọ̄/ (Allen, 1987, p. 72). In Pompeii a word usually written λάσϑη, *lást*^{h}*ē*, is instead written λάσφη, *lásp*^{h}*ē*, indicating a confusion reminiscent of the cross-linguistic tendency to confuse [f] and [θ] and suggesting fricativization of the aspirated plosives had already begun (Allen, 1987). Similarly, Latin spelling of *cum* 'with' (which did not have an inherited labialized velar) as *quom* (a historically correct spelling of the conjunction *cum* 'when', which did have an inherited labialized velar) in a 2nd century BCE epitaph suggests phonetic merger of /kw/ and /k/ in this environment (Clackson & Horrocks, 2007, p. 149).

Evidence from other languages Evidence can also be gleaned from languages which were related to or in contact with Greek. Borrowed words, and the representations of these Greek words in the borrowing languages, occasionally provide evidence for Greek pronunciation at the time of borrowing. Evidence for a dental articulation of Greek <τ> in the first and second centuries BCE comes partly from coins minted by Greek-speaking rulers in Bactria and India, in which Greek <τ> was rendered with a Prakrit dental rather than the contrastive alveolar /ṭ/ (Allen, 1987, pp. 16-17). Similarly, the Greek personal name *P*^{h}*ilippos* was rendered in early Latin writing as *Pilipus*, indicating that at the time of borrowing Greek <φ> indicated a sound closer to that of Latin /p/ than to any other sound for which Latin had a grapheme at the time; Latin had a grapheme <f> for /f/, so the use of <p> for Greek <φ> suggests that <φ> represented a stop, not a fricative as it did in later Greek. Later borrowings were transcribed into Latin with *ph*, either reflecting the development of aspirated stops in certain positions in Latin (Allen, 1989) or indicating a conventional realization that the sounds in Greek could be more accurately represented by the use of two graphemes in Latin; the fact that <ph> is not used earlier to transcribe Greek <φ> is due to this digraph not having been invented, rather than a lack of aspiration in Greek (Probert, 2010, p. 86). Regardless of its origin, the new spelling helps to confirm the presence of aspiration as a contrastive feature in Greek. Evidence that an aspiration contrast was retained for some speakers long after the classical period, rather than fricativizing the inherited aspirated plosives, comes from Greek words and alphabetic symbols borrowed into other languages; the Coptic writing system devised in the 3rd century CE, and the Armenian and Georgian alphabets of the 5th century CE, all used <ϑ, φ, χ> to represent their aspirated plosives (Allen, 1987, p. 23). Transcriptions of Greek names on Indian coins help to support the identification of <β, δ, γ >as voiced stops rather than voiced aspirates (Sturtevant, 1940; Allen, 1987).

Metre Many surviving ancient texts are written in verse which had strict metrical requirements, providing a source of evidence on pronunciation. For example, the Latin alphabet did not generally distinguish between long and short vowels, but in instances in which a vowel appears in a heavy syllable—but is not followed by more than one consonant, which would also generally make the syllable heavy—then the vowel must have been long (Allen, 1989, p. 64–65). This and similar concepts are relevant to the labialized velar question for reasons discussed in §2.2.4.

Combining evidence It is generally necessary to combine multiple types of evidence before making conclusions about ancient pronunciation. In combination, these sources of evidence allow some reliable conclusions regarding pronunciation of Greek in the historical period and the types of contrast that were phonemically meaningful.

2.2 Reconstruction of labialized velars

2.2.1 Evidence for a three-way velar contrast in PIE

A labialized velar stop series in PIE is reconstructed on the basis of comparative evidence from IE daughter languages, both from modern languages which provide direct phonetic evidence and from dead languages whose phonetic details are ascertained using the methods described in §2.1. Much of this evidence points to a series of voiceless unaspirated, voiced unaspirated, and voiced aspirated labialized velars ($*k^w$, $*g^w$, $*g^{wh}$). Traditionally, three separate series of stops characterized by the presence of velarity are reconstructed for PIE: the plain velars ($*k$, $*g$, $*g^h$); the palatal velars ($*\acute{k}$, $*\acute{g}$, $*\acute{g}^h$) and the labialized velars ($*k^w$, $*g^w$, $*g^{wh}$). Labialized velars are reconstructed based on the presence of both labiality and velarity in different daughter languages, and surviving labialized velars in some languages. Correspondence sets supporting the reconstruction of a three-way PIE contrast in the velar stop series are presented in Table 2.1. In Luvian, distinct outcomes are observed for the plain, palatal, and labialized velars. In Sanskrit, a *satem* language, the plain velars and labialized velars show convergent outcomes, and the palatal velars remained distinct. In Greek, Latin, and Gothic—*centum* languages—the plain and palatal velars merged and the labialized velars remained distinct.

Table 2.1: Sound correspondences illustrating the three-way reconstruction of PIE velars. Some conditioned developments are not included in the table. See e.g. Fortson (2010) for an overview of the developments in each language.

PIE	Luvian	Sanskrit	Greek	Latin	Gothic
*_k_	/k/	/k, c/	/k/	/k/	/h/
*_leuk-_ ‘light’		_rócate_	_leuk-_	_lūc-_	_liuh-_
*_kes-_ ‘comb’	_kiš-_		_késkeon_		_heord_
*_ḱ_	/z/	/ś/	/k/	/k/	/h/
*_ḱerd-_ ‘heart’	_zārt-_	_śrád-_	_kard-_	_cord-_	_hairto_
*k^{w}	/ku/	/k, c/	/p, t, k/	/kw/	/hw/
*$k^{w}i$-, $k^{w}o$- ‘who’	_kui-_	_kás_	_tis_	_quis_	_hwas_
*_g_	/k/	/g, j/	/g/	/g/	/k/
*_gras-_ ‘eat’		_grás-_	_grā-_	_grāmen_	
*_ǵ_	/k/	*_ȷ́_ > /j/	/g/	/g/	/k/
*_ǵenu-_ ‘knee’	_genu-_	_jānu_	_gónu_	_genū_	_kniu_
*g^{w}	/w/	/g, j/	/b, d, g/	/g, gw, w/	/kw/
*$g^{w}em$- ‘come’		_gam-_	_bain-_	uen-	_qim-_
*$g^{w}en$- ‘woman’	_wāna-_	_jáni-_	_gun-_		_qino_
*g^{h}	/k/	/g^{h}, h/	/k^{h}/	/f, g, h/	/g/
*$steig^{h}$- ‘climb’		_ati-stígh-_	_steík_^{h}-		_steigan_
*$ǵ^{h}$	/k/	/h/	/k^{h}/	/f, g, h/	/g/
*_ǵheu-_ ‘pour’	_kūtt-_	_hūyáte_	k^{h}_é(w)-_	_fundō_	_giutan_
*g^{wh}	/w/	/g^{h}, h/	/p^{h}, t^{h}, k^{h}/	/f, gu, w, g/	/gw, w/
*$g^{wh}en$- ‘slay’		_hán-_	t^{h}_ein-_	_dē-fen-_	

2.2.2 Evidence for labiality and velarity in the daughter languages

This study aims to use phonetic evidence to draw conclusions regarding the most plausible mechanisms of sound changes involving labialized velars. In order to do this accurately, it is necessary to first characterize the phonetic details of the labialized velars, to the extent possible given the challenges inherent to the discipline; this enables more effective comparisons with the modern languages used as proxies in Chapters 3 and 4. Whilst the *satem* languages do not provide evidence of labiality as the labialized velars merged with the plain velars, evidence from multiple *centum* languages is presented in the following section to provide as full a description as possible of the phonetics of PIE labialized velars.

Greek Within Greek, labialized velars underwent conditioned sound changes described in Table 1.1. The Greek sound changes are discussed in detail in §2.5.

2.2.3 Anatolian

Luvian and Lycian may be the only languages preserving evidence that distinct series of plain and palatalized velars (as well as labialized velars) were inherited by the *centum* languages, and are therefore to be reconstructed for PIE: the apparent distinction is not simply an artefact of the internal history of the *satem* group (as has been argued by e.g. Sihler 1995, p. 151 ff.). In addition to preserving evidence of this contrast between plain and palatalized velars, Lycian appears to show evidence of an intermediate stage in development between PIE $*k^w$ and a coronal reflex before front vowels.

Three-way velar distinction Reflexes of PIE voiceless unaspirated labialized velars are generally represented by *ku* syllabic signs in Hittite, indicating that they retained labialized velar articulation; see Examples 2.1-2.2. The picture for the voiced series is more complex due to later sound changes: inherited $*g^w$ became /w/ in all positions in Luvian and is also represented by the spelling *ku* in Hittite, suggesting labiality and velarity in both languages (Example 2.3). The voiced aspirated stops merged with the voiced unaspirated stops in proto-Anatolian (Fortson, 2010, p. 156), meaning there is no distinction between /gw/ and /gwh/ in Hittite or Luvian.

(2.1) Hitt. *nekuz* 'of evening' from PIE $*nek^wt$- 'night' (Kloekhorst, 2014).

(2.2) Hitt. *kuis*, relative pronoun, from PIE $*k^wis$, relative pronoun (Kimball, 1999, p. 86).

(2.3) Hitt. *kuinna-* and Cuneiform Luvian *wāna-*, 'woman', from PIE $*g^wen$- 'woman' (Melchert, 1994, p. 120, 254).

The original PIE three-way distinction between series of plain, palatalized, and labialized velars was at least partly preserved in the Anatolian languages (Melchert 2012; Melchert 1987 Fortson 2010, p. 156; see Sihler 1995, p. 154 for an alternative viewpoint). As outlined in Table 2.1, Luvian shows evidence of divergent developments of the three types of velar stop in the voiceless unaspirated series: the reflex of PIE $*k$ is /k/; the reflex of $*\acute{k}$ is /z/ in front vowel environments and before approximants, and /k/ elsewhere (Melchert, 2012); and the reflex of $*k^w$ is /ku/. Examples 2.4-2.7 provide evidence of this three-way velar distinction in Cuneiform Luvian for the voiceless unaspirated series. Melchert (2012) argues that Luvian shows evidence of development as a *centum* language, merging $*\acute{k}$ and $*k$ in front vowel environments, as can be seen from Examples 2.5 and 2.6.

(2.4) CLuv. *-ku* 'and', from PIE $*k^we$ 'and' (Melchert, 1994, p. 95, 252).

(2.5) CLuv. *zārt-* 'heart', from PIE $*\acute{k}rt$- 'heart' (Melchert 1994, p. 252; Kloekhorst 2014).

(2.6) CLuv. *kattawatnalli* 'spiteful', from PIE **ḱot*, cf. Skt. *śátru-* 'enemy', Gk. *kótos* 'spiteful' (Melchert, 2012).

(2.7) CLuv. *kiš-* 'comb', from PIE **kes-* 'comb' (Kloekhorst, 2014).

The picture for the voiced velars is less clear, partly due to a lack of clear examples in the extant corpus and partly due to secondary phonological changes in Luvian and Lycian, including medial loss of voiced stops (Melchert, 1994, 2012).

Lycian is a poorly attested branch of Luvian, written in a local alphabet adapted from the Greek (Melchert, 1994, p. 39). Lycian also apparently shows a three-way distinction between the different velars; Example 2.8 shows a coronal outcome of *k^w, Example 2.9 shows a velar outcome of **k*, and Example 2.10 shows a sibilant outcome of **ḱ*.

(2.8) Lyc. *ti-*, relative pronoun, from PIE *$k^w i$-, relative pronoun (Neumann, 2007).

(2.9) Lyc. *tukedr(i)-* 'statue', from PIE **twek-* 'body', cf. Skt. *tvác-* 'skin' (Melchert, 1994, p. 284).

(2.10) Lyc. *si-* 'lie' from PIE **ḱei*, cf. CLuv. *ziyari* 'lies' (Fortson 2010, p. 148, 173; Kloekhorst 2014).

Development o f*k^w to /t/ **in Lycian** The Lycian alphabet has four symbols which appear to be some kind of velar stop: <k, q, x, K>; and two symbols which apparently transcribe dental sounds: <t, τ>. The phonetic values of these letters are difficult to determine (Kloekhorst, 2008; Melchert, 1994). The two signs <t>and <τ> appear to alternate with one another (Kloekhorst, 2008, p. 125); generally <t> corresponds to Greek τ (Kloekhorst, 2008), indicating a dental quality, and appears in positions where a labialized velar is reconstructed before a front vowel as well as in other positions where a coronal **t* is reconstructed, suggesting convergence between the two. However in some words, <τ> is used in positions where a labialized velar is reconstructed before **e*, as in Examples 2.11 and 2.12 (although all of the examples also have spellings with <t>, Neumann 2007). This suggests that the sign may represent an intermediate stage between *k^w and /t/. A discussion of the possible phonetic value of the Lycian <τ> is provided in §5.3.1.

(2.11) Lyc. *τeτeris* and *teteris* 'four', from PIE *k^w*etwor-* 'four' (but the meaning is disputed; see e.g. Neumann 2007, p. 398).

(2.12) Lyc. *τezi* and *tezi* 'sarcophagus' from PIE *k^w*iē-ti-* 'rest' (Pedersen, 1945, p. 50), cf. Lat. *quiēs*, Av. *š(ii)a* 'rest'.

2.2.4 Italic languages

Latin retained the inherited PIE labialized velar, but neighbouring Italic languages show labial outcomes.

Latin Labialized velars survived in Latin and were transcribed *qu* with evidence indicating that *q* had a voiceless velar quality and *u* was a labialized velar approximant.

(2.13) *quis* 'who' from PIE *k^w*is*, cf. Gk. *tis* 'who'.

(2.14) *linqu-* 'leave' from PIE **link*^{w}-, cf. Gk. *leip-* 'leave'.

(2.15) -*que* 'and' from PIE *-k^w*e*, cf. Gk. *te* 'and'.

The velar quality of *q* is accepted partly on the basis of evidence from later Romance languages, which confirm the velarity of the sounds represented by these letters. In early Latin inscriptions three letters were used to write velar sounds depending on the vowel environment, a tradition retained from the adaptation of alphabets in Italy from the Greeks, via the Etruscans: <C> before front vowels, <K> before /a/, and <Q> before /o, u/ (Wachter, 1987, p. 15). However, whatever difference in pronunciation <C, K, Q> represented was not contrastive in Latin and, with the exception of <QU> sequences, the use of <C> was generalized (Wachter, 1987).

Evidence for the pronunciation of Latin /u/ comes from several sources. For example, Nigidius Figulus, writing in the 1st century BCE, provides a physiological description which refers to protrusion of the lips during the articulation of <u> (Gellius, x, 4, 4; Allen 1989, p. 41). Additionally, Cicero describes an unfortunate confusion between *Cauneas*, a placename, and *caue ne eas*, 'don't go!', which led to the downfall of Marcus Crassus (Cicero *De Divitatione*, ii.84; Allen 1989). Other descriptions (Allen, 1989; Sturtevant, 1940) suggest roundedness of the <qu> sequence; for example, Marius Victorinus, writing in the 4th century CE, says that <k> and <q> do not differ in place of articulation with respect to the throat, but do differ in the protrusion of the lips: this presumably refers to the use of <q> before <u> to transcribe a labialized velar:

> Nonnihil tamen enterest utra earum prior sit, [c] seu *q* sive *k*. Quarum utramque exprimi faucibus, alteram distento, alteram producto rictu manifestum est.
>
> *Nevertheless it makes some difference which of them precedes,* q *or* k*. It is clear that both of them are pronounced in the throat, one with the mouth wide open, the other with the mouth-opening drawn forward.*
>
> Marius Victorinus, GL vi. 34. 1-3; trans. Sturtevant (1940)

Osco-Umbrian Latin retained the inherited PIE labialized velar, but neighbouring related languages show labial outcomes:

(2.16) PIE *k^wa-*, interrog. pronoun, giving Oscan *paam*, cf. Lat. *quam*.

(2.17) PIE *k^wetwor-* 'four', giving Umbrian *petur-* 'four-'.

(2.18) PIE *pek^w-* 'cook', giving Osc.-Umbr. *popina* 'eating-house' subsequently borrowed into Latin, cf. Lat. *coquīna* 'cookery'.

Sardinian Some Romance languages changed the Latin labialized velars to labials, as in Sardinian.

(2.19) PIE *k^wetwor-* 'four', giving Sardinian *battoro* 'four'.

(2.20) Lat. *aqua* 'water', giving Sardinian *abba* 'water'.

Romanian Labialized velars developed to labials before /a/ during the development from Latin to modern Romanian (Examples 2.21–2.23). There is one example of a labial outcome before /i/ (Example 2.24). Generally, the outcome of inherited /kw/ before /i/ in Romanian is /tʃ/, but this is also the outcome of inherited /k/ so it is not possible to determine whether /kw/ and /k/ each independently coronalized in this position or whether /kw/ underwent delabialization first and merged with /k/ before coronalizing (Examples 2.25–2.26).

(2.21) Lat. *aqua* 'water', giving Rom. *apă* 'water'.

(2.22) Lat. *lingua* 'tongue', giving Rom. *limba* 'tongue'.

(2.23) Lat. *quattuor* 'four', giving Rom. *patru* 'four'.

(2.24) Lat. *liquidare*, giving Rom. *lepid-* .

(2.25) Lat. *quem*, rel. pronoun, giving Rom. *cine* /tʃine/, rel. pronoun.

(2.26) Lat. *cena* 'dinner', giving Rom. *cină* /tʃina/.

2.2.5 Celtic

At an early stage in the history of Celtic languages, inherited PIE $*g^w$ developed to /b/ (Examples 2.27–2.29).

(2.27) PIE $*g^w en$ 'woman', giving OIr. *ben* and Wel. *ben*.

(2.28) PIE $*penk^w e$ 'five', giving Wel. *pimp*.

(2.29) Proto-Celtic $*k^w ids\underset{\frown}{i}\bar{o}m\bar{\imath}$ giving Gaulish *pissíiumí* (Fortson, 2010, p. 315).

2.2.6 Germanic

Labialized velars also persisted in some Germanic languages, including English. The following series of sound changes between PIE and proto-Germanic affected—along with the other inherited stops—all three labialized velars, and are often referred to as Grimm's Law:

(2.30) $*k^w > *h^w$, e.g. PIE $*k^w od$, cf. Goth. *hwa-*, Eng. *what, who*.

(2.31) $*g^w > *k^w$, e.g. PIE *$*g^w en$-*, cf. Goth. *qēns* ‘woman’, Eng. *queen.*

(2.32) $*g^{wh} > g^w$, e.g. PIE $*seng^{wh}$-, cf. Goth. *siggwan* ‘sing’, and $*g^{wh} > w$, e.g. *snaiws* from PIE $*sneig^{wh}$- ’snow’.

These outcomes are attested in extinct Germanic languages, such as Gothic, and contemporary languages, including some varieties of English in which the reflex of PIE labialized velars is articulated as a voiceless labialized velar fricative [ʍ] or the ‘standard’ variant voiced labialized velar approximant [w].

2.3 Proto-Greek labialized velars: stops or clusters?

That labialized velar sounds existed in some branches of PIE is, therefore, confirmed by strong evidence. However, there is less certain evidence to indicate how the PIE labialized velars stood in relation to clusters of a velar followed by a labial approximant. The question of what exactly constitutes a labialized velar stop, what constitutes a labialized velar cluster, and how they may differ phonetically and phonemically in different languages, is addressed in Chapter 3.

There are few opportunities to compare outcomes of labialized velars to the outcomes of clusters of a velar and a labialized velar approximant, due to the relative scarcity of examples of clusters; the majority of roots that do have clusters feature voiceless, unaspirated velars with few examples of the voiced or voiced aspirated stops, leaving something of an imbalance in the available evidence. Understanding the nature of labialized velar stops and their segmentation in Ancient Greek is essential to drawing conclusions based on typological data. Comparative evidence from within Ancient Greek and from related languages will be assessed in order to address this question.

2.3.1 *Satem* vs. *centum* languages

The apparent existence in PIE of contrasting sequences of **k* or **ḱ* followed by **w* has been said to indicate some phonetic or phonemic difference between the single labialized velar and the cluster. This evidence comes almost exclusively from those languages which retained a distinction between palatal and plain velars; see examples 2.33-2.39.

(2.33) **eḱwos* 'horse', gives Skt. *áśvas* and Hieroglyphic Luvian *asuwa* (Melchert 1989; the word is discussed further in §2.3.2).

(2.34) **ḱwon-* 'dog', gives Lith. šuõ and Hieroglyphic Luvian *zuwan-* (Melchert, 1989).

(2.35) **ǵhwēr-* 'wild animal', gives Lith. zvěrǐ.

(2.36) possibly **kwend^h r-*, 'plant', giving Lat. *combrētum*, a type of plant, and Lith. *švendr-* 'reed', O.Ice. *svǫnn* 'Angelica silvestris'. The etymology is proposed by Pokorny (1959) and rejected by de Vaan (2008), after Heiermeier (1980).

(2.37) possibly **ǵ^{h}woig^w-* 'shine', giving Gk. *p^{h}oibos* 'shining' and Lith. *zvaigzde* 'star' (Pokorny, 1959), but this has not been explored further.

(2.38) probably not **ǵ^{h}welg-* 'glance'; see Table 2.2.

(2.39) probably not **ǵ^{h}wel-* 'bend'; see Table 2.2.

In *satem* languages the inherited labialized velars show different, non-labialized outcomes, as shown by examples 2.40-2.44.

(2.40) **k^{w}ol-* 'turn', gives OCS *kolo*.

(2.41) **k^{w}o-*, interrogative pronoun, gives Skt. *kás*.

(2.42) **sneig^{wh}* 'snow', gives Russ. *sneg*.

(2.43) **ǵhwēr-* 'wild animal', gives Lith. zvěrǐ.

(2.44) **k^{w}is*, interrogative pronoun, gives Cuneiform Luvian *kuiš* (Melchert, 1989).

In contrast, relatively abundant examples of an apparent merger of **kw* and *k^w in *centum* daughter languages are available. Evidence from these languages is discussed in the following sections.

2.3.2 Greek

Comparisons with the inherited labialized velar approximant Developments of inherited **w* in Greek—namely its loss, at different times, in all positions and in almost all dialects—differ from the developments of inherited labialized velars, lending support to the hypothesis that labialized velars did not behave as clusters of a velar stop plus a labialized velar approximant in proto-Greek. Actual clusters of non-velar stops plus **w* developed in very different ways, with the outcomes apparently never including labiality (see examples 2.45-2.52, from Sihler 1995). This indicates that—at least at the stage at which **w* disappeared from Greek—the labialization of the labialized velar differed in treatment from the labialization of stop + *w* clusters.

(2.45) Gk. *dís* 'twice' from PIE **dwis*, cf. Lat. *bis*, Skt. *dvís* 'twice'.

(2.46) Gk. *sé* acc. sg. 'you' (but also Doric τρέ for (presumably) *twé*, Hesych.), cf. Skt. *tvam* acc.sg 'you', from PIE **tu-* or **tw-* 'you'.

(2.47) Gk. *sákos* '[oxhide] shield', cf. Skt. *tvak-* 'cow hide', Hitt. *tuekka-* 'body' (possibly a Semitic loanword, cf. Akk. *saqqu* and Hebr. *śaq*, Beekes 2010).

(2.48) Att. *kṓra* 'girl' from earlier attested *korwa* 'girl', inc. Myc. *ko-wa* (KN Ag 1654, PY Aa 62 etc.), from P.Gr. *u*-stem *korwā* which comes from an original PIE root **ḱerh₃* 'grow'.

(2.49) Att. *kalós*, Ion. *kālós* 'beautiful' from earlier **kalwos* 'beautiful'.

(2.50) Att. *mónos*, Ion. *mōnos*, from **monwos* 'alone'.

(2.51) Att. *hólos*, Ion. *hōlos* 'all' from **solwos*, cf. Skt. *sárva-*.

(2.52) PIE **k^wetwores* ‘four’ > Ion. *tésseres*, Att. *téttares*, Boeot. *pettares*, Lesb. *pésures*, Dor. *tétores*.

Monophonemic outcomes in Greek Additional evidence that labialized velars were stops rather than clusters in proto-Greek is provided by the monophonemic outcomes (§2.5.2). There is some evidence that there was a distinction between reflexes of /kw/ and/k^w/; this may have been a difference in length, as the cluster in PIE **eḱwos* ‘horse’ (Example 2.33) shows the same labial development in its Greek outcome *híppos* as labialized velars elsewhere, only doubled in length. However, this is perhaps the only Greek word that provides good evidence of such a divergent outcome, and it is widely regarded as problematic: there is no good account of the change in vowel quality or the acquisition of an initial glottal fricative (Fortson 2010, p. 186; Hooker 1980). The word for ‘horse’ is also attested in Mycenaean Greek. Mycenaean is the earliest known variety of Greek, and in most positions in which labialized velars have been reconstructed for early Greek they are transcribed in Mycenaean using the <*qV*> series of syllabic signs. The syllabic script — designed for an unknown pre-Greek language — does not provide conclusive evidence of whether this series was meant to represent a single stop, a cluster onset, or a geminate; however, the spelling of ‘horse’ as <*i-qo*> in Mycenaean (discussed in detail in §2.5.2) demonstrates that the <*qV*> series was thought suitable to represent the outcome of an inherited PIE **ḱw* sequence. There is one other possible example of an inherited cluster represented by the apparently monophonemic *-qV-* series in Mycenaean Greek, namely <*qe-ri-jo*>, personal name (Example 2.53). Ventris & Chadwick (1973) identified the name as *K^hwēríōn*. However, the subsequently identified example of a genitive <*qe-ri-jo-jo*> suggests that the nominative singular ending cannot be *-ōn*, leading to other possible explanations, such as that the name originates from the same root as *Tēlemak^{h}os* and therefore has a labialized velar, rather than a cluster (Aura Jorro, 1993).

(2.53) Myc.<*qe-ri-jo*> personal name, cf. later Gk. *T^hēríōn*, personal name, possibly from PIE **ǵ^{h}wer-* (Aura Jorro, 1993).

Thus there is little evidence to suggest that labialized velars and clusters of /kw/ were distinct in Mycenaean, due to apparently convergent orthography. However, subsequently formed combinations of /k/ and /w/ whose separate components were still productive clearly differed in articulation from the outcome of that merger, prompting the use of *ku-wV* spellings rather than the *qV* series. The quality of Mycenaean /w/ is impossible to determine for sure, but it could be the case that it had diverged sufficiently from the inherited PIE labialized velar approximant that it no longer resembled the secondary articulation of the inherited labialized velar stops, to the extent that an alternative orthography was justified (cf. development to [v] in Tsakonian, e.g. *βάννϵ* 'sheep' from proto-Greek **wam-no-*). Vis (2011) supports this interpretation with analysis of the synchronic behaviour of Mycenaean /w/ in clusters and in alternation with *u*. On the other hand, the reverse could be the case: namely, that labialized velars had developed in such a way that their labialized velarity no longer resembled that of the labialized velar approximants. It is worth noting that the Mycenaean *wV* series still retained enough rounding for it to be chosen to represent a glide between rounded vowels (e.g. Myc. *du-wo* 'two', cf. later Greek *duō*). Either way, it seems that the Mycenaean writing system implies two differing articulations for inherited and 'new' clusters of /k/ and /w/ following the merger of the inherited sounds. Exactly what the nature of those articulations was, however, remains uncertain.

The small number of proposed PIE roots that feature labial/velar clusters complicates the analysis. Proposed correspondence sets and reconstructions—for those roots which have reflexes in Greek—are provided in Table 2.2. It is relevant here to note that, without cognates from at least one *satem* language, it may not be possible to determine whether a root should be reconstructed with a labialized velar or with a labial/velar cluster, due to merger in many daughter languages. As can be seen from Table 2.2, those correspondences which have been proposed are not often supported by good evidence; many have been challenged or have alternative explanations. Perhaps the best supported reconstruction is that of PIE *g^h*wer-* 'wild animal',

Table 2.2: Proposed PIE roots with labial/velar clusters that have possible Greek reflexes.

Root	Greek	Possible Cognates	Source
$*$*kwres-* 'wood'	*prīnos* 'oak'	OIr. *crann* 'spinney', Rus. *chvórost* 'deadwood', Slov. *hrást* 'oak'; alternatively, Slav. *brin-* 'larch'	Pokorny (1959); Beekes (2010) and Frisk (1972) connect to Slav. *brin-* as a non-IE loanword
$*$*kwēt-* 'winnow'	*pētea* 'grain'	Lat. *quat-* 'shake', MIr. *cāith*; connection to Toch. *kät* refuted (Adams, 2014)	Pokorny (1959) (doubtful)
$*\acute{k}weh_2$-	Att./Ion./Arc. *pāma*, Boeot. *ppám-* 'possession'	None	Rix (2001); Beekes (2010)
$\acute{g}^h$*wel-* 'bend'	p^h*olk-* 'bow-legged', p^h*ál-* 'helmet band (?)'	Skt. *hvár-* 'go crooked', Av. *zbar-* 'crooked', Lat. *fal-* 'cheat', Lith. *žval-* 'agile'	Pokorny (1959); Beekes (2010) designates substrate; Frisk (1972) gives no etymology
$*\acute{g}^h$*wer-* 'wild animal'	t^h*er* 'wild animal'	Lat. *fer-* 'wild', Skt. *hvár-*, Av. *zbar-*, Lith. *žvèr-* 'beast'	Pokorny (1959); Rix (2001); Beekes (2010); Frisk (1972)
$*\acute{g}^h$*welg-* 'glance'	t^h*élg-* 'bewitch'	Lith. *žvelg-* 'look at', Skt. *hvár-* 'go obliquely'	Numerous hypotheses from Beekes (2010); Frisk (1972); Pokorny (1959); Rix (2001) rejects Baltic connection on phonological basis
$*\acute{g}^h$*woig*^{w}- 'shine'	p^h*oīb-os* 'gleaming'	OCS *dzvěz-* 'star'	Pokorny (1959); not generally accepted Beekes (2010)

which has explicable reflexes in Greek, Latin, Sanskrit, Avestan, and numerous Balto-Slavic languages. In all the *centum* languages in which this word appears, its initial cluster shows the same outcome as the equivalent inherited labialized velar.

2.3.3 Anatolian

Proto-Anatolian preserved (in certain positions) a distinction between plain, palatalized, and labialized velars which continued into Luvian and Lycian; the partial merger of plain and palatal velars suggests Anatolian should be categorised as *centum* (Melchert 2012; §2.2.3). However, $*\acute{k}w$ sequences and labialized velars show contrasting outcomes in Luvian and Lycian, as in *satem* languages; see Examples 2.54-2.55 (Melchert, 2012).

(2.54) HLuv. *azuwa-* and Lyc. *esb-* 'horse', both from PIE **eḱwos*.

(2.55) HLuv. *zuwan(i)-* 'dog', from PIE **ḱwon-*.

Although the syllabic cuneiform script makes it difficult to draw conclusions about stop clusters in Hittite, alternation of *ku* and *uk* in positions where labialized velars would be expected may indicate a monophonemic identity for inherited /k^w/, e.g. *ekuzi* and *eukzi* for 'he drinks', *tarkuzi* and *tarukzi* 'he dances' (Kloekhorst, 2014). Other evidence is seen in exceptions to sound changes affecting the sequence /uw/. Hittite changed inherited **w* to /m/ after **u* word-internally, e.g. *warnu(m)meni*, composed of *warnu-* 'see' and *-weni*, 1st pl. verbal ending, and possibly also *tarumaki-* 'woodpecker' from *taru-* 'wood' and *waki-* 'bite' (Hoffner & Melchert, 2008, vol. 1, p. 44). When verbal inflection would have created the conditioning environment in which this change would be expected to affect a sequence *ku*, the outcome remained *ku*, indicating a possible phonetic difference between /k^w/ and *ku*. For example, Hitt. *akueni*, 1.pl.pret.act. from PIE *$eg^{w(h)}$- 'drink' (cf. Lat. *aqua* 'water', Gk. $n\bar{e}p^h$- 'be sober, not drink') is attested rather than ***ekmeni* (Hoffner & Melchert, 2008; Kloekhorst, 2014; Kim, 2000).

2.3.4 Tocharian

Tocharian—a *centum* language whose geographical location challenged the early theory that there was an isogloss in PIE which depended upon whether plain velars merged with labialized velars, as in 'satem' languages, or with palatalized velars, as in 'centum' languages—may provide some evidence for divergent developments. The outcome of **eḱwos* in Tocharian B is *yakwe*, and Adams connects Toch. B *kwäs* 'mourn' to Skt. *śvásiti* 'groan' via PIE *$\acute{k}wesh_x$-. Meanwhile, the single stop PIE /k^w/ had velar outcomes following the collapse of all PIE velars to voiceless unaspirated *k* in proto-Tocharian, e.g. Toch. A *ak* from PIE *h_3ekk^w- 'eye', Toch. A *ko* from PIE *g^wou- 'cow', and Toch. AB *ts*. However, other correspondences seem to prove almost the reverse. Toch. A *saku* and B *sekwe* 'pus' apparently reflect PIE

sok^w 'sap', cf. Gk. *opós* 'sap' (Adams 2014; Lith. *sakaĩ* 'resin' appears to confirm the PIE labialized velar). Thus the situation in Tocharian seems somewhat complex, and the data scarce, meaning it is difficult to draw many meaningful conclusions on phonemic or phonetic distinctions between /kw/ and/k^w/ in these languages or in PIE.

2.3.5 Latin

In Latin, there is some evidence to indicate that the sound represented by the letter sequence *qu* (and therefore presumably also *gu*, which only survived after a nasal) was recognised by speakers as one segment, rather than a cluster, but there are some apparent exceptions which complicate the analysis. First, in poetry, syllables in which a short vowel is followed by *qu* were generally scanned as short, as if *qu* represented a single stop, rather than long as would be the case if it was a cluster; for other clusters, e.g. *tr* and *tl*, the syllable is only sometimes short. Grammarians, too, appear to describe lip-rounding as coinciding with the onset of the velar constriction, indicating a simultaneous articulation (Allen, 1989, p. 16). In addition to this, the *u* articulation in *qu* and *gu* did not develop into a fricative, unlike plain *u*. In fact, no other clusters of any onset plosive and *u* are attested in Latin, a fact that may lend support to the theory that *qu* and *gu* were special, single segments—or, alternatively, provide a conditioning environment which might explain the resistance to fricative developments. The question is discussed at length in Devine & Stephens (1977, p. 85–99).

Generally, developments of PIE labialized velars and velar/labial clusters do not differ from one another in the Italic languages (Sihler, 1995, p. 180); for example **mag-welō* 'prefer' gives **māwolō* and eventually *mālō*. Similary, PIE **$mre\acute{g}^h w$-i-* 'short' gives Lat. *brevis* (Sihler, 1995; de Vaan, 2008). A link between Lat. *cāseus* and OCS *kvas-* 'sourdough', Goth. ƕaþ- 'froth' from PIE *kwat-* 'ferment', has been proposed (Pokorny, 1959); however, this would be an isolated case of the loss of *w* in Latin in this position, and the etymology is regarded as uncertain (de Vaan, 2008).

2.3.6 Germanic

That labialized velars and clusters fell together in the pre-history of Germanic languages appears clear, e.g. in the voiceless unaspirated series, PIE **kwat-* 'ferment' gives Goth. ƕaþō 'froth' (the reconstruction of the cluster is supported by e.g. OCS *kvas-* 'sourdough'), cf. ƕi- from **k*

Evidence for early Germanic languages is found in Gothic; the sound represented by the letter generally transcribed as ƕ resulted from *k^w by Grimm's law and probably sounded similar to the [ʍ] observed in some modern dialects of English. The Gothic letter *q* likely represented/k^w/ from PIE *g^w, again via Grimm's law; evidence includes its use to transcribe Biblical names otherwise spelt with Greek letters κου- Marchand (1973). The representation of these sounds by a single character may indicate that they were perceived by speakers to be single sounds rather than clusters, although such an argument seems less valid after considering the situation as described above for Latin, where two letters are used to describe a sound which seems to have been phonemically unified. Marchand argues that /ƕ/ and */q/* were clusters, but his argument rests solely on the structuralist assumption that, because there are other clusters of stops + /w/, there must therefore also be a cluster of /k/ + /w/ and /h/ + /w/; no other evidence, from syllabification rules or otherwise, is provided.

2.3.7 Celtic

The usual outcomes for PIE labialized velars in Irish are exemplified by Examples 2.56-2.59, in which labiality is lost but velarity retained. In other Celtic languages, labial outcomes are the norm.

(2.56) **k^wetr*- 'four' gives Ir. *cethair*, MWel. *pedwar*, Corn. *peswar*, Bret. *pewar*, (Matasović, 2009).

(2.57) **g^wenh$_2$* 'woman' gives Ir. *ben* and poetic *bé* 'woman', OWel. *ben* 'woman', OBret. *ban-doiuis* 'woman-goddess', OCorn. *benen* 'bride' (Matasović, 2009; Jasanoff, 1989).

(2.58) **neig^w*- 'wash' gives Ir. *nigid* (Pokorny, 1959).

(2.59) **sneig^{wh}*- 'snow' gives Ir. *snigit* (Pokorny, 1959).

There is some evidence for the retention of labiality in velar labial clusters in Irish, i.e. they show a different outcome from the labialized velars, as in Examples 2.60-2.61. This suggests that they remained distinct in this branch.

(2.60) **kwend-ro-*, a type of plant, giving OIr. *cuinneog* 'Angelica silvestris' (Pokorny, 1959), cf. Lith. *šveñdrai*, a type of reed, Dan. *quander*, OIce. *huǫnn* 'Angelica silvestris'.

(2.61) **ḱwon-* 'dog' giving OIr. *cú* with retention of labiality (Jasanoff, 1989).

However, other Celtic languages show evidence of convergent developments which indicate a merger of labialized velars and labial velar clusters. In Examples 2.62-2.64 the outcomes of velar labial clusters have labiality in these languages, like the outcomes of labialized velars shown previously in Examples 2.56-2.59 (although note the non-labial Irish outcome in Example 2.64).

(2.62) **skwerb(^{h})*- 'stick, pierce', giving Corn. and Bret. *spern* 'spinae', cf. Lith. *skverbiú* 'to prick' (Pokorny, 1959)

(2.63) **gwos-to-* 'branch' giving OIr. *bos*, OWel. *bos*, MBret. *boz*, cf. Alb. *gjethe* 'leaf', MHG *quast* 'branch' (cf. OCS *gvozdъ* 'nail' from the same root with a different suffix, **gwos-d^{h}o-*, Matasović 2009).

(2.64) **sk^{h}uoi-* 'needle, spike', giving OIr. scē 'hawthorn', Wel. *ysbyddad* 'hawthorn', Corn. *spethes* 'bramble', Bret. *spezad* 'gooseberry', cf. Russ. *xvojá* 'conifer needles' and Lith. *skujà* 'fir-needle' (Matasović, 2009).

2.3.8 Conclusions

The only languages which do seem to unequivocally show a difference in outcomes between labialized velars and clusters of a velar and a labial are those which fall into the *satem* group, and thus would presumably not have had a labialized velar with which **kw*, **gw* etc. could have fallen together. With the exception of Anatolian, it is rare that the *centum* languages allow us to detect a distinction between the reflexes of $*\acute{k}w$, $*\acute{g}w$, $*\acute{g}^h w$ and $*k^w$, $*g^w$, $*g^{wh}$; presumably this is due to the phonetic similarity between the sounds and the lack of palatalization as a cue to differentiate the cluster. In any case, it appears that the phonetic and phonological boundaries between the cluster and the stop may have been somewhat blurred in the *centum* daughter languages; it is entirely possible that merger could have taken place independently in various language families or groups. It is likely that typological analysis of the articulatory and acoustic relationships between clusters and double-articulation stops in modern languages, and the phonetic cues that differentiate them, could shed light on the questions raised in this section.

2.4 The conditioning vowels

2.4.1 Vowels in PIE

The direction of development of the inherited labialized velars in Greek was determined partly by the quality of the following vowel (see Table 1.1). Early Greek vowels are generally thought to have continued the quality of the PIE vowels relatively faithfully (Allen, 1987; Sturtevant, 1940; Sihler, 1995), and in fact the reconstructed qualities of the Indo-European vowels are often based upon the proposed qualities of those in Greek (Fortson, 2010). These are described in Examples 2.65-2.69.

(2.65) **a*, e.g. Gk. *hál-*, Lat. *sal-*, Arm. *ał*, OIr. *sal-* ‘salt’.

(2.66) **e*, e.g. Gk. *hép-*, Lat. *sequ-*, OIr. *sech-*, Lith. *sekù* ‘follow’.

(2.67) **i*, e.g. Gk. *tis*, Lat. *quis*, Skt. *cit*, OCS *či-to*, relative pronoun.

(2.68) **o*, e.g. Gk. *bówes*, Lat. *bouēs*, OIr. *boin*, Arm. *kou*, OCS *gov-* 'cow(s)'.

(2.69) **u*, e.g. Gk. *zugón*, Skt. *yugám*, Lat. *iugum*, Hitt. *i-ú-kán*, Goth. *juk* 'yoke'.

The status of **a* in PIE is somewhat controversial. It has been argued that the vowel did not exist at all in early PIE (Lubotsky, 1989) but instead resulted separately in the daughter languages from the vocalization of **h_2* or the coalescence of **h_2e*. Alternative views suggest that it did exist but occurred infrequently in early PIE, and multiplied in later PIE and the daughter languages through processes such as those mentioned above and the vocalization of syllabic sonorants, as in Greek and Latin (Mayrhofer, 1986, Vol. 1, p. 170; Sihler, 1995; Meier-Brügger, 2003); thus the comparative evidence for the quality of *a* in Greek is perhaps less useful than circumstantial evidence from within Greek and from its interactions with other languages. The vocalic phonemes as they are generally described for Greek are given in Examples 2.73-2.77. IPA suggestions are given for reference only and are not meant to constitute an absolute statement on the phonetic quality of PIE vowels.

Long vowels were apparently rare in early PIE. Many PIE reconstructions containing long vowels are the result of contractions of a short vowel adjacent to a laryngeal, as in the case of Example 2.71. Since there is sparse direct evidence for laryngeals surviving in any daughter branch except Anatolian, there is little to gain from making a distinction between inherited long vowels and long vowels resulting from laryngeals (Fortson, 2010, p. 60-61). Cognate sets favouring the reconstruction of long vowels in PIE are presented in Examples 2.70-2.72 (cognates from Fortson 2010, p. 60).

(2.70) **ē*, e.g. Skt. *rā́ṭ*, Lat. *rīx*, Gaulish *-rīx*, 'king'.

(2.71) **ā*, e.g. Skt. *mātár-*, Dor. *mā́tēr*, Lat. *māter*, OIr. *máthair*, OCS *mater-*, 'mother', from PIE **māter* from earlier PIE **meh₂ter* 'mother'.

(2.72) $*\bar{o}$, e.g. Skt. *svásā*, Lat. *sorōr*, OIr. *siur*, 'sister'.

2.4.2 Vowels in Greek

The vowels of proto-Greek are generally reconstructed as outlined in Examples 2.73-2.77; the long vowels corresponded in quality (Allen, 1987).

(2.73) /a/, an open central unrounded vowel, e.g. [ä].

(2.74) /e/, a short mid front unrounded vowel, e.g. [e].

(2.75) /i/, a short close front unrounded vowel, e.g. [i].

(2.76) /o/, a back close-mid rounded vowel, e.g. [o].

(2.77) /u/, a back rounded vowel, e.g. [u].

In Attic-Ionic varieties of Greek in the alphabetic period the sound represented by *υ* was fronted, sounding something more like [y]. It is not possible to state definitively at what time that change occurred, but it seems to have been quite early; inscriptions of the 6th century B.C.E occasionally show <αο> and <εο> for diphthongs /au/ and /eu/, suggesting that perhaps the <υ> grapheme at that time was not suited to transcribe a back vowel (Allen, 1987, p. 63). However, the back round vowel persisted in other dialects: Boeotian writers used <ου> to transcribe their /u/, rather than adopting the Attic convention of using <υ>, when they adopted the Attic alphabet in the 4th century BCE. There is no evidence for fronting in the other older daughter languages (or even in other pre-Koine Greek dialects) indicating that the change was limited to some dialects of Greek and that the *u* sound in PIE and in early Greek was probably not particularly fronted.

Long vowels are not included in Table 1.1 due to the difficulty of tracing the complex developments of long vowels in Greek and the various Greek dialects. In the alphabetic period Greek had seven long vowels; these are transcribed in this thesis as *ā*, *ē*, *ei*, *ī*, *ō*, *ou*, *ū*. The sounds *ē* and *ō* (η and ο) reflect inherited PIE long vowels, whilst the digraphs *ei* and *ou* (ει and ου) represent long vowels resulting from secondary processes in Greek, including compensatory lengthening, contraction of adjacent short vowels, and monophthongization of inherited diphthongs probably completed by the late 5th century BCE. The vowel inventory of many Ancient Greek dialects, including Attic, had more long vowels than short counterparts, and the long vowels differed in their quality: *ei* was probably a close-mid vowel and *ē* an open-mid vowel and *ou* was closer than *ō* (Allen, 1987). Thompson (2006) argues that the inherited long vowels became more open as a result of structural pressure from the new, more close long vowels. Despite these changes, the outcomes of the labialized velars after the long vowels appear to match up to those of the equivalently articulated short vowels, although examples are very scarce and not particularly well-supported by comparative evidence (Examples 2.78-2.82).

(2.78) **iēg^w-ā* 'youth' giving Dor. *hḗbā*, cf. Lith. *jegá* (Frisk, 1972; Beekes, 2010).

(2.79) Proposed **k^{w}ēl-* 'far' giving Att. *tēle*, Aeol. *pḗlui* (loc)., Boeot. *Peile-* (component in personal names) (Pokorny, 1959); there is a possible but unclear connection with Att. *pálai* 'furthest', semantically similar but with a short vowel, and *télos* 'goal, ending', from a proposed root **telh$_2$-* 'lift, carry' (Frisk, 1972; Beekes, 2010). The reconstruction is complicated by Myc. *te-re-ta* 'master of ceremonies' (cf. the sense of 'executive function, office' linked to Gk. *télos*) (Chantraine, 1980; Beekes, 2010, p. 1103).

(2.80) Possibly **kwēt* 'winnow', see Table 2.2; the connection to Gk. *pḗtea* is doubtful.

(2.81) **k^{w}āli-*, *tāli-*, giving Att.-Ion. *pēlíkos*, Dor. *tāl-* 'how big', Att.-Ion. *hēl'ikos* 'how big', cf. Lat. *quālis* 'what kind, *tālis* 'such'.

(2.82) **ǵ^{h}uēr*; see Table 2.2 for cognates; the short vowel in Lat. *fer-* 'wild' is most likely the result of pretonic shortening before a resonant in Latin, after Schrijver's interpretation of Dybo's Law (Schrijver, 1991, pp. 334–357).

It is difficult to determine if all the changes in consonant phonotactics that co-occur with compensatory lengthening in Greek had yet taken place in Mycenaean Greek (these changes are generally thought to constitute several successive developments:

loss of **s* in certain conditions; simplification of inherited word-internal resonant and apical fricative clusters and contraction of adjacent short vowels; and simplication of secondary word-internal resonant and apical fricative clusters, and of word-final **ns* Colvin 2007; Thompson 2006, p. 10) . The first change of this type affected clusters of a resonant and /s/ in non-final syllables. The fact that verbal endings in -*onsi* resulting from the change of -*onti* to -*onsi* are not affected indicates that the cluster simplification change postdates the assibilation of the verbal ending; this implies that cluster simplification must already have taken place in Mycenaean, since Mycenaean only ever has the -*onsi* variant (Thompson, 2006, p. 83-85). Thompson continues by arguing that rather than being retained or being reduced to single segments with compensatory lengthening of the preceding vowel, Mycenaean **rs* clusters had undergone gemination without compensatory lengthening, as is observed in Thessalian and Lesbian dialects of later alphabetic Greek. To show this he presents as an example the spelling <*a-ke-ra*$_2$*-te*> for *agerrantes*, from an earlier participle **ager-s-antes*, cf. later Gk. *ageírō* from **ageriō* 'gather'. There are no examples of this type of gemination co-occurring with compensatory lengthening in later dialects. Thus the evidence indicates that compensatory lengthening had not taken place in Mycenaean (assuming that word-final stop loss occurs before lengthening) and thus Mycenaean cannot offer any meaningful evidence to inform conclusions about the relative chronology of labialized velar changes and long vowel development, beyond the statement that the first wave of cluster simplification had likely already taken place in Mycenaean. Additionally, since Myceanean does not appear to have been the direct ancestor of any Greek dialects (see §2.5.2) conclusions based on Mycenaean are not always obviously applicable to later attested Greek.

2.5 Development of labialized velars in Greek

2.5.1 Overview

An introductory picture of the Greek changes was given in Table 1.1, and the developments are discussed in greater depth here. Due to the inconsistency of outcomes in the different dialects, reflexes specific to each dialect are presented in §2.5.2-2.5.7.

2.5.2 Mycenaean Greek

Mycenaean Greek, although the oldest attested dialect of Myceanean Greek, seems not to have been the ancestor of any later attested dialects. This is demonstrated by innovations found in Mycenaean but not in all other dialects of Greek; for example, /o/ vocalization of syllabic nasals adjacent to labial consonants as in <*pe-mo*> *spermo*, from **spermn* 'seed', cf. classical Gk. *spérma* (although <*pe-ma*> is also attested, Horrocks 2010, p. 20–21), and possibly accusative plurals in *-es* (Probert, 2008, p. 139, n. 16). The exact relationship of Mycenaean to the other Greek dialects is complex and controversial, and the arguments will not be dealt with in depth here. Mycenaean has some features in common with East Greek dialects (see §2.5.3); it shares innovations with Attic-Ionic and Arcado-Cypriot, such as the change of inherited verbal endings *-ti* to *-si* and the use of <*o-te*> 'when' rather than West Greek *hóka* Horrocks (2010); Colvin (2010). Within the East Greek subgroup, Myceanaean appears to share additional similarities with the Arcadian and Cypriot dialects. For example these dialects feature, amongst other innovations, /o/ vocalization of inherited syllabic liquids and the use of the dative rather than the genitive after prepositions meaning 'out of' or 'from' (Colvin, 2010; Thompson, 2010). However, Mycenaean also has innovations that are not shared by Arcadian or Cypriot, such as the extension of the thematic infinitive *-hen* into the athematic conjugation, suggesting it is closely related to the two later dialects but not their common ancestor.

Mycenaean Greek uses a series of syllabic signs to represent sounds in positions where labialized velars have been reconstructed for early Greek using the comparative method. These are generally transcribed as *qa, qe, qi, qo* (there was no *qu* symbol). Examples 2.83-2.91 show a selection of Mycenaean attested labialized velars.

(2.83) <*-qe*> (KN Le 641, PY Aq 64, etc.), *k^{w}e*, from PIE *-*k^{w}e* ‘and’, cf. Lat. -*que*.

(2.84) <*qa-ra-to-ro*> (PY Ta 709.2), *sk^{w}alat^{h}ron* ‘oven rake’, cf. later Gk. *spalat^{h}ron* ‘oven rake’ (Beekes, 2010; Aura Jorro, 1993).

(2.85) <*qe-to-ro-po-pi*> (PY Ae 27.a, etc.), *k^{w}etropop^{h}i* ‘four-footed’, from PIE **k^{w}etwores* ‘four’ and **pod*- ‘foot’.

(2.86) <*qi-ri-ja-to*> (KN Ai 1037.2), *k^{w}riato*, [aor.] ‘buy’, from PIE **k^{w}reih$_2$*- ‘buy’, cf. later Greek *príast^{h}ai* ‘to buy’, Skt. *kríta* ‘bought, OIr. críth ‘purchase’, Toch.B. *karyor*, Toch.A. *kuryar* ’purchase, trade’ (Beekes, 2010).

(2.87) <*qa-ra$_2$*> (PY An 192.16), *K^{w}alians* ‘Pallas’, cf. later Gk *Pállās, -antos*, divine name (Aura Jorro, 1993).

(2.88) <*to-ro-qe-jo-me-no*> (PY Eq 213.1), *trok^{w}eiomenos*, possibly ‘walking around’ although the meaning is not certain (Aura Jorro, 1993) from PIE **trk^w*- ‘twist’; cf. later Gk. *trépō* and Hom. *tropéō* ‘turn’, Lat. *torqueō* ‘twist’, Hitt. *tarku*- ‘dance’, Skt. *niṣ-ṭarkyà* ‘what can be unscrewed’.

(2.89) <*qo-u-ko-ro-jo*> (PY Ea 781), *g^{w}oukoloio* ‘cow-herd’, from PIE **g^{w}ouk^{w}ol*-, with dissimilation of **k^w* adjacent to **u* (see §2.6.1). <*a-pi-qo-to*> (PY Ta 709, etc.), *amp^{h}ig^{w}otos* possibly ‘accessible from all sides’, from PIE **amp^{h}i* ‘around’ and **g^{w}m*- ‘go’ (Bartoněk, 2003; Ventris & Chadwick, 1973, p. 138)

(2.90) <*to-ro-qa*> (KN Fh 358), *t^{h}rok^{wh}ā* ‘nourishment’, cf. later Gk. *trep^{h}ō* ‘feed’ (Bartoněk, 2003, p. 129).

(2.91) <*po-ru-qo-ta*> (PY Jn 845.2 etc.), *Poluk^{wh}ontās*, cf. Hom. *Polup^{h}ontēs*, personal name (although there are other possibilities for the second component, e.g. -*g^{w}ōtās*, cf. later Gk. *Polubṓtēs*, Aura Jorro 1993).

The exact phonetic realisation of the *qV* series is, however, not known. The fact that they occur in positions for which labialized velars are posited by comparative reconstruction does not mean that the sounds they represented were necessarily still labialized velar in nature. Evidence of alternative spellings for the same word—one spelling with the <*qV*> series, and another using a different spelling—would help in settling this question, and it has been suggested that such examples exist. The Linear B syllabary had signs for a small and not comprehensive group of complex onset syllables, including those with labial-velar elements to the second constituent: <*dwe*>, <*dwo*>, <*nwa*>, <*twe*>, and <*two*> e.g. Example 2.92.

(2.92) <*dwo*> (PY Ub 1315 etc.), *dwo* 'two', cf. later Gk *dúo*, Skt. *dvā́*, Lat. *duo*, Goth. *twai* etc.

However, some combinations of stop and labial-velar approximant in syllable onset position are represented with two symbols rather than the equivalent single symbol, as in Example 2.93.

(2.93) <*ma-ra-tu-wo*> (MY Ge 602.2), *marat^h^won* 'fennel' (Aura Jorro, 1993, p. 424–425).

There are several examples which demonstrate that the two symbol spelling could alternate with the single symbol spelling. For instance, in Examples 2.94 and 2.95 the sequence /nwa/ is represented by Linear B syllabograms <*nu-wa*> and <*nwa*> on different tablets. The same may also be true of Example 2.96. The Mycenaean identification of *Enualios* is not considered secure because the /ū/ is long in later Greek, which—unless it is the result of metrical lengthening in order to make the word fit into hexameter—would require unparalleled phonological changes Morpurgo Davies 2009, p. 109 and indicates a different syllabic structure in which the sequence /nua/ is disyllabic, rather than constituting a single syllable with a complex onset. On the other hand, there is an inscriptional example of *Enuálios* with a short /u/ (Lyr. Adesp. 108). In Examples 2.97 and 2.100, the sequence /dwo/ is represented by the syllabograms <*du-wo*> and <*dwo*> on different tablets.

(2.94) Myc. <*qa-nu-wa-so*> (KN As 1516.17) and <*qa-nwa-so*> (KN D1 943.B), both interpreted as variant spellings of nom. masc. personal name *K^wanúas(s)os* (Aura Jorro 1993, p. 184, Morpurgo Davies 2009, p. 109).

(2.95) Myc. <*pe-ru-si-nwa*> (MY Oe 111.51) and <*pe-ru-si-nu-wa*> (PY Ub 1316), *perusinwós* 'last year', from PIE **per* 'other side' and **uet* 'year', cf. later Gk. *pérusi(n)* and Dor. *péruti(n)* 'last year', Arm. *heru* 'last year', MHG *vert* 'last year', OIr. *ónn-urid* 'last year' (Beekes, 2010; Aura Jorro, 1993).

(2.96) Possibly Myc. <*e-nu-wa-ri-jo*> (KN V 52.2) and <*e-nwa-ri-jo*> (PY An 724.1) for the divine name *Enuálios* (Aura Jorro, 1993, p. 221).

(2.97) Myc. <*te-mi-de-we-te*> (PY Sa 1266.a) and <*te-mi-dwe-te* (PY So 4433.a) for *termidwéntes* 'reaching to the feet', cf. Hom. *termióessa* 'fringed' (Aura Jorro, 1993, vol. 2, p. 328).

(2.98) <*o-two-we-o*> (PY An 261.43), <*o-tu-wo-we*> (PY Jn 658.7), <*o-to-wo-we-i*> (Vn 851.9), *Ort^hwōwēs*, personal name (Aura Jorro, 1993, vol. 2, p. 55).

(2.99) <*wi-do-wo-i-jo*> (PY An 5, Ae 344.22) and <*wi-du-wo-i-jo*> (PY Hn 415.3) and <*wi-dwo-i-jo*> (PY Eb 1186.a), *Widwóhios*, personal name (Aura Jorro, 1993, vol. 2, p. 428)

(2.100) Myc. <*du-wo-jo*> (Py Jn 750.12) and <*dwo-jo*> (KN X 8126), *Dwoios*, personal name (Aura Jorro, 1993, vol. 1, pp. 198–200).

There are examples of sequences spelt <*kV-wV*>. These are generally examples of sequences with the syllabification *CVwV*, i.e. sequences of two syllables in which the second has the labial-velar approximant as an onset. For the purposes of this discussion, relevant examples include /u/ as the first vowel, as in Example 2.101 and confirm that this spelling was available to scribes to represent sequences such as *kuwV*.

(2.101) <*ku-wa-no*> (PY Ta 642), *kuwanos* 'lapis lazuli', cf. later Gk. *kúanos* with regular loss of intervocalic **w*.

Evidence of the <*qV*> series alternating with spellings in <*kV-wV*>, as observed for other clusters in Examples 2.94-2.100, would provide evidence for an articulation with labiality and velarity. However, no secure examples are available. One example that has been proposed is the word for 'horse' (the reflex of PIE **eḱwos*, later Greek *híppos* 'horse', Dor. *híkkos* 'horse'), which appears in the Mycenaean tablets as <*i-qo*>. This in itself suggests that the inherited cluster **ḱw* sounded sufficiently similar to *k^w at this stage to be represented by the same series of syllabograms in Mycenaean. It was proposed by (Palmer, 1966) that *i-ku-wo-i-pi* (KN V 280) represents a dual form of <*i-qo*>. If this were the case it would suggest that <*qo*> and <*ku-wo*> were both acceptable spellings for the outcome of an inherited sequence **ḱwo*, presumed to be similar in realisation to *$k^w o$, and thus the <*qo*> sign at least had labialization and velarization. However, this etymology rests on shaky foundations: it relies upon assertions about possible Mycenaean religious beliefs that cannot be proven and thus should not be accepted without additional evidence, which has not yet been found, and the context is problematic due to the difficulty of interpreting the rest of the inscription on this tablet.

Although there are no examples of <*qV*> alternating with <*kV-wV*> spellings for the same word, there are examples in which a velar stop appears adjacent to a labial-velar approximant as the result of morphological operations, e.g. the addition of the participle suffix *-woha* (as in Example 2.102). Where this happens, the cluster is always transcribed with <*kV-wV*> and never with the <*qV*> notation (Szemerényi, 1966). This suggests either that the perceptual status of these morphological boundaries was sufficient to prevent their falling together orthographically with the inherited labialized velars, or that they sounded different from the labialized velars, or both.

(2.102) <*te-tu-ko-wo-a*> (KN L 871.b), *tetuk*^{h}*woha* 'made', cf. later Gk. *teuk*^{h}- 'make' plus active perfect participle suffix *-woha*.

(2.103) <*pa-ra-ku-we*> (PY Ta 714.1) and <*pa-ra-ke-we*> (PY Ta 642.1), possibly *parakwei* or *barakwei* 'of the colour *baraku*'; cf. Hesych. *barakís* (Aura Jorro, 1993). The word seems to be a *u*-stem noun *paraku-* or *baraku-* appearing here with the dative ending *-ei* (Aura Jorro, 1993).

The possibility that labialized velars sounded sufficiently different from velar and labial clusters that it was not possible to transcribe them in the form <*Cv-wV*>, unlike with other clusters, indicates that they may have already developed towards their eventual Greek outcomes in Mycenaean Greek. Example 2.104 shows orthographical confusion between a sound /tw/ and a sound /kw/ in a front vowel environment.

(2.104) Myc. <*o-da-ke-we-ta*> (KN So 4446.131), <*o-da-ku-we-ta*> (KN L 870), <*o-da-tu-we-ta*> (KN So 894.4), and <*o-da-twe-ta*> (KN So 4430.b), for *odatwenta* 'provided with teeth'; presumably a compound of **odnt-* 'tooth' giving *odat-* with /a/ vocalization of the syllabic nasal, and the adjectival suffix *-went-*.

This may be an indication that the labialized velar had already begun to develop towards something similar to /t/. Szemerényi (1966) pursues this argument further by connecting <*qi-nwa-so*> (Example 2.105) to <*ti-nwa-si-jo*> (Example 2.106), and proposes that inherited *k^w had progressed to /tš^w/ before front vowels in Mycenaean Greek causing possible spelling confusion between <*qi*> and <*ti*>. The word <*ti-nwa-si-jo*> itself is linked to later Greek *t*^{h}*ís* 'heap, dune' (Heubeck, 1963) but does not have a plausible Indo-European etymology (Frisk, 1972).

(2.105) <*qi-nwa-so*> (KN Dc 1515.B), personal name.

(2.106) <*ti-nwa-si-jo*> (PY Jo 438.21) 'rich in sand'.

The dialects most closely related to Mycenaean, Arcadian and Cypriot, show coronal outcomes in front vowel environments (/s/ in Cypriot, see §2.5.5 for details on

Arcadian developments), suggesting that a coronal-type development would not be unexpected for Mycenaean Greek. On the other hand, an alternative explanation for the alternations in Example 2.104 is proposed by Aura Jorro (1993), who suggests dissimilation of /kw/ under the influence of the neighbouring /t/ in the adjectival suffix, as well as confusion with the similar word *odaks* 'with the teeth'. There is also no precedent for a labialized velar to result from dissimilation from /t/ in Mycenaean Greek, unlike in the case of the labials in Examples 2.107 and 2.108. Additionally, in the case of the labials and labialized velars, the stops share labial articulation as a possible prompt to dissimilation, whilst /kw/ and /t/ are less obvious candidates for such a change. The connection between Examples 2.105 and 2.106 is complicated by the lack of a good etymology for *t^{h}ís*, as there is no way to be certain the word actually had a labialized velar in the first place. Thus these examples do not provide convincing evidence either way for the phonetics of Mycenaean labialized velars in front vowel environments.

Some Mycenaean words in which inherited labialized velars or labial-velar sounds appear to alternate with the Linear B <*pV*> series have been discussed as possible indications that some labialized velars had begun to show labial outcomes. Examples 2.107 and 2.108 demonstrate this.

(2.107) <*i-po-po-qo-i*> (PY Fn 79), *hippoporg^{w}oihi* 'horse-keepers', cf. Myc. <*i-qo*>'horse'.

(2.108) <*qe-re-qo-ta*> (PY En 659.1), *K^{w}ēlek^{wh}ontās*, and <*pe-re-qo-ta*> (PY An 192.12 etc.) *Pēlek^{wh}ontās*, personal name; cf. later Gk. *Tēlep^{h}óntās* (Ventris & Chadwick, 1973).

However, these alternations can be explained by the presence of another labial or labialized velar stop occurring later in the same word, prompting anticipatory dissimilation of the labialized velar stop. There are no examples of <*pV*> appearing where comparative reconstruction would indicate we should expect <*qV*> in which there is not another labial or labialized velar nearby.

Thus there is some evidence suggesting that labialized velars before front vowels may have had a more palatal articulation in Mycenaean Greek, but it is rather sparse and not entirely convincing. The only examples pointing towards a non-velar labial articulation of $<qV>$ in Mycenaean can all be explained by processes of dissimilation near a second labialized velar. However, it seems clear that the sounds represented by the $<qV>$ series in Linear B had some other phonetic or phonological status than clusters of other obstruents and /w/, as shown by the orthographical flexibility applied to clusters (including, apparently, combinations of /k/ and /w/ resulting from morphological operations) but not to reflexes of inherited labialized velars.

2.5.3 The Greek dialects

Following the collapse of the literate Mycenaean civilization in the 12th century BCE and the subsequent 'Dark Ages', the advent of alphabetic writing by the 8th century BCE resulted in numerous alphabetic inscriptions and texts which reflected the geographical diversity of spoken Greek at the time. Traditionally, four major dialect groups are proposed for Greek at this point in time: Attic-Ionic, Aeolic, Arcado-Cypriot, and West Greek. A further connection is often made between Attic-Ionic and Arcado-Cypriot, forming a higher level grouping called East Greek (e.g. Colvin 2010, p. 205). Phonological, morphological, and syntactic differences between the dialects are many and there is a great deal of scholarly literature regarding linguistic relationships and population movement in this period; however, as this thesis is concerned with labialized velars the focus of the current discussion will be limited to the ways in which the classical dialects differed in their treatment of those sounds.

The situation is of course not as straightforward as this type of dialect classification would suggest; like any language, the Greek dialects were affected by contact and population movement over many centuries, making it impossible to really separate the dialects. Approaches to categorization of the dialects are summarized in e.g. Colvin 2010 and Horrocks 2010, p. 9–40.

2.5.4 Attic-Ionic

Attic was the dialect spoken in Attica, a region in Greece that contains Athens, in the classical period. Ionic was spoken on the west coast of Asia Minor (modern day Turkey), on the large islands of Chios, Samos, and Euboea, and on other smaller islands in the Aegean Sea. Some phonological details related to their treatment of labial-velar sounds are given here. Attic and Ionic both lost the inherited labial-velar approximant before alphabetic times (Example 2.109) in contrast with most other dialects in which it survived in inscriptions (Sihler, 1995, p. 182).

(2.109) PIE **woiḱ-* 'house', gives Att.-Ion. *oíkos*, Myc. <*wo-i-ko-de*> *woikonde*, 'homeward', Thess., Boeot., West Gk., Arc. *woíkos* 'house', cf. Lat. *vīc-* 'house', Skt. *váśa* 'house'.

For Attic, the outcomes of labialized velars are as described in Table 1.1. For Ionic, developments are the same except that there is an apparent alternation between forms with *po-* and forms with an alternative pronominal stem *ko-*, both from PIE *k^wo. This is generally *po-* in other dialects. These forms are almost entirely confined to literary Ionic, however, rather than Ionic inscriptions (Buck, 1955, p. 63).

(2.110) PIE *k^wo-, pronominal stem, gives e.g. Ion. *kót*^{h}*en* 'from somewhere', cf. *pót*^{h}*en* in other dialects; Ion. *kou* 'to where', cf. *poi* 'to where'; Ion. *kóteros* 'which of two', cf. *póteros* 'which of two'.

Suggestions for how a velar stop might become generalized throughout the pronominal paradigm are provided in §2.5.6.

It is unclear why the word *obelós* 'metal bar used as a coin or weight' has a labial in Attic-Ionic, where a coronal would be expected; in West Greek and Arcadian the cognate is *odelós*. The alternation of /d/ and /b/ indicates an original labialized velar, but there is no obvious PIE etymology (Chantraine, 1980, p. 772). The word has been connected to PIE $*g^w el$- 'throw', which has other Greek reflexes including *bállō* 'throw' and *bélos* 'missile', but there is no good explanation for the initial /o/ of *obelós* and *odelós*. A Hesychian gloss *odolkaí*, with coronal outcome after a back vowel, complicates the issue further. Beekes (2010) believes the word is not of Indo-European origin.

2.5.5 Arcado-Cypriot

The dialects of Arcadia and Cyprus were linked by modern historical linguistic scholarship, and not by ancient writers (Colvin, 2010). This may be because in the classical period the two groups were separated by considerable geographical distance, presumably the result of population movement following a period of closer development. They share some traits with Mycenaean, such as development of the third person verbal ending *-ti* to *-si* and *o*-vocalization of PIE syllabic resonants (see §2.5.2), suggesting a close relationship between the groups. Reflexes of labialized velars in Arcadian and Cypriot show some kind of affricated or assibilated treatment before front vowels. Since most evidence postdates the period in which the two dialects developed concurrently, the following sections deal with Arcadian and Cypriot separately.

Arcadian

In inscriptions of the fourth and third centuries, Arcadian shows developments similar to West Greek and Attic-Ionic, e.g. *odelós* 'coin, weight' (cf. Att. *obolós*, Thess. *obellós*). However, in earlier inscriptions the situation seems less clear; the reflexes of labialized velars were often still orthographically (and therefore presumably also phonemically) distinct from coronals, and appear to have been represented by several different orthographies; see Examples 2.111-2.116. In several examples, Σ is used either preceding or following a T (Dubois, 1986, p. 67).

(2.111) ζ′ for *zd'*, elided enclitic 'and', IPArk 20, at Pheneos; from PIE **k^w^e*.

(2.112) τζετρακατιαι for *tzdetrakatiai*, IG V.2.159, at Tegea; from PIE **k^w^etr-* 'four'.

(2.113) ζτηραιον for *zdtēraion*, unidentified, at Cleitor/Lusai, but cf. Hesych. *zdeirón*) 'spotted'.

(2.114) Arc. ζερεχρα *zdéret^h^ra*, given as a gloss for Att. *bárat^h^ron* 'pit' (Chantraine 1980, p. 164; Strabo *Geographica*, 8.8.4.5; Hesychius).

(2.115) In compounds, inscriptional *-déllō*, cf. Att. *bállō* 'throw' (Schwyzer 1939, p. 295; Buck 1955, p. 58).

(2.116) Arc. *ózdis* 'whoever' from PIE **k^w^is* 'who' (Colvin, 2007, p. 33).

In a fifth century inscription from Mantinea, IG V. 2. 262, the reflexes of labialized velars before front vowels are transcribed with a special character of uncertain phonological value: И. This is the only example of the reflexes of a labialized velar being represented by a special symbol in alphabetic Greek (Duhoux, 2007). The character appears ten times in the Mantinean evidence:

(2.117) OИEOI, line 14: *oИeoi*, cf. Ion. *'otéōi*, Att. *hóitini*.

(2.118) AΠΥИEΔOMIN[..], line 19: *apuИedomin[..]*, cf. Aeol. *apudedómen[..]*.

(2.119) EIΔAΛΛA ИI[.], line 23: *eidallaИi[.]*, various interpretations.

(2.120) ИIΣ, lines 25, 26: *Иis*, cf. Cyp. *si-se*, Att. *tís*, Thess. *kis*.

(2.121) EIИand EIИE, lines 26 (x2), 27, 28, 31: *eiИ*, *eiИe*, cf. Att. *eíte*.

The same symbol is also used in an Arcadian bronze *lex sacra*. The text is currently published by (Heinrichs, 2015) but the readings provided here are from a personal communication with James Clackson. Interestingly, in this text И appears in combination with ς to represent the outcome of PIE **tw*:

(2.122) ИΕΣИΑΡΟ, line 11: *ИesИaro* 'four', from PIE **k^wetwor-* 'four'.

(2.123) ΟИΕΛΟ, ΟИΕΛΟΝ, line 13, 19: *oИel-*, cf. Att.-Ion. *obolós*, Arc. *odelós*, Thess. *obellós* 'obol'.

The problem is in establishing the phonological value assigned to the letter И. Other graphemes similar in shape occur elsewhere, but do not seem to provide useful evidence in determining the phonological value of the Arcadian И; an identically shaped character in Pamphylian appears as a variant of digamma, whilst a Lycian character /\/\ is likely to denote a velar or palatal consonant, but is unrelated to the Arcadian letter. The identity of Carian И is uncertain, while И at Melos and Selinus is a variant of beta (Brixhe, 1996, p. 6). The character Τ was already in use to represent a labialized velar reflex before Ι, as we can see from other Arcadian inscriptions of the 5th century (e.g. <ΤΙΣ>, IG V 2, 429,5; Lillo 1988). This would indicate that whatever sound the inscriber meant to represent with И probably had some dental or apical quality to it.

The use of Z to represent some kind of affricate in early inscriptions of some Greek dialects (such as the Cretan reflex of an inherited **ty*, Allen 1987) and the proposed routes of articulatory development outlined above have led some to interpret И as a voiceless affricate, /tʃ/ (Colvin 2007, p. 338; Colvin 2010, p. 208; Panayatou 2007, pp. 418–419). Agreement is not unanimous, however, and Chadwick (1988, p. 58) discounts /tʃ/ as a possible phonological value of И on the basis of Cypriot evidence; he follows Buck in identifying sibilant developments of a labialized velar before a front vowel as a shared feature of Arcadian and Cypriot, but rejects the suggestion that the Cypriot syllabary would transcribe a sound /tʃ/ with <*s*> as in e.g. <*si-se*> (corresponding to Att. *tís*), so a phonological value of /ʃ/ is more likely for И. On the other hand, the loss of the labialized velars may have occurred late enough in the history of Arcado-Cypriot that divergent developments could be attested in Arcadian and Cypriot. The labialized velars were retained in Mycenaean, which is closely related to Arcadian and Cypriot, suggesting they could also have been retained in both dialects following the geographical separation of Arcadian and Cypriot; this would provide an opportunity for similar results to be achieved via different developmental mechanisms in each branch. However, as argued in §2.5.2, Mycenaean is not the direct ancestor of these other two dialects, and it cannot be assumed that labialized velars persisted in Arcadian or Cypriot at the time Mycenaean is attested. An explanation in which each dialect retained labialized velars and separately innovated sibilant reflexes—the only Greek dialects to do so—seems like a suspicious coincidence, and it is more likely that some period of concurrent development took place and that labialized velars had developed into some other sound before the separation of Arcadian and Cypriot.

The use of И in line 23 of the Mantinea inscription presents further difficulties as there are a number of possible readings. Guarducci 1937 suggests that И could here be representative of /ks/, based on an interpretation of ΑΛΛΑИΙ[.] as equivalent to *állaksis* 'exchange'. This can be rejected partly on the basis that a different sign, +, is used consistently in other sixth century inscriptions found at Mantinea in positions where Ξ would be expected (Lillo, 1988, p. 86). Thus the reading of Comparetti (1914), followed by Buck (1955), of the sequence as some form of the relative pronoun $*k^w is$ is more plausible.

The Mantinean inscription has led to several philological arguments proposing incremental articulatory developments of PIE $*k^w$ to Gk. /t/, via whichever sound the Arcadian symbol was intended to represent. These are discussed in §5.3.

Cypriot

The Cypriot dialect was written using the Cypriot syllabary, a Cypro-Minoan writing system, from the 8^{th} to the 3^{rd} centuries BCE. Like Linear B, the Cypriot script was not entirely well suited to writing Greek; for example, it did not distinguish contrasts in voicing or aspiration (Colvin, 2007, p. 20). Cypriot uses syllabograms representing <*si*> for the reflex of PIE *$k^w i$, e.g. <*si-se*> *sis*, 'who' from PIE *$k^w is$ (the second vowel is purely orthographic, Woodard 2010, p. 33).

2.5.6 Aeolic

The Aeolic group is a label given to the dialects Thessalian, Lesbian, and Boeotian. Aeolic dialects are generally said to show labial outcomes of labialized velars even before a front vowel. Szemerényi (1966) puts forward an alternative argument in which there are no regular labial outcomes of labialized velars in Aeolic, but rather individual exceptional explanations for each apparent example; whilst this concept has not been widely accepted, there are some exceptions to the Aeolic labial outcomes (Colvin, 2010, p. 210). Examples of labial outcomes are given below (Examples 2.124-2.129, from Buck 1955, p. 58).

(2.124) PIE *$k^w etwor$- 'four', gives Att.-Ion. *tessares*, Lesb. *péssures*, Boeot. *péttares* 'four'.

(2.125) PIE *$penk^w e$ 'five', gives Att.-Ion. *pente*, Lesb. *pémpe* and Thess. *pémpe* 'five'.

(2.126) *IE *$k^w ei$- 'punish, avenge', gives Att.-Ion. *teìsai* Thess. *peìsai*, Boeot. *potapopisátō*.

(2.127) PIE *$k^w el$- 'far', gives Att.-Ion. *tḗle* 'far', Lesb. *pḗlui* 'far', Boeot. *Peile-strotídas*, personal name.

(2.128) PIE *$g^w el$- 'want', gives West Gk. *dḗlomai*, Thess. *béllomai*, Boeot. *beílomai* 'want'

(2.129) Att. *T^h^essalía* ‘Thessaly’, Thess. *Pett^h^alía* ‘Thessaly’, from an earlier *K^{wh}*etyal-*.

In some instances where a labialized velar is reconstructed before a front vowel, labiality is lost, and the outcome looks similar to those of other dialects:

(2.130) PIE *$k^w e$* ‘and’, gives Att.-Ion. *te* ‘and’; in adverbs, Lesb. *-ta*, Dor. *ka*.

(2.131) PIE *$k^w i$*, ‘who’, gives Att.-Ion. *tis*, Thess. *kis*.

(2.132) PIE *k^w*ei-* ‘observe, honour’, gives Att.-Ion. *tīmḗ* ‘honour, value’, Dor. *timā́* ‘honour, value’, Boeot. *tim-* ‘honour’ (Myc. <*qe-ja-me-no*> confirms the labialized velar).

(2.133) PIE **sm̥-g^w^elb^h^-* ‘sibling’, gives Att.-Ion., Lesb., and Boeot. *adelp^h^ós* ‘brother’.

(2.134) PIE *g^{wh}*erm-* ‘warm’ gives Lesb. *t^h^érm-* ‘warm’.

Example 2.135 may provide additional evidence for coronal outcomes in Aeolic and is presented as such by Stephens & Woodard (1986), but the etymology is disputed. Beekes (2010) rejects the connection of Aeol. *déra* ‘neck’ to PIE *g^w*er-* ‘devour’ partly on the basis of the absence of a labial reflex.

(2.135) *déra* ‘neck, collar’; Frisk (1972) connects it to the PIE root *g^w*er-* ‘devour’ (as seen in e.g. Gk. *bibrōskō* ‘devour’), cf. Skt. *grīvā-* and Rus. *gríva* ‘neck’, Lat. *grĩva* ‘mouth (of a river)’; alternatively, Beekes (2010) connects it to Gk. *deirás* ‘mountain ridge’, of unknown origin.

There is no obvious phonological conditioning environment for the alternation of /p/ and /t/ in front vowel environments in the Aeolic dialects. Whilst some of these vocabulary items could be explained by borrowing from non-Aeolic dialects, it seems less likely (although not impossible) that words such as *te* ‘and’ and *tis* ‘who’ would have been borrowed.

Thessalian *kis*

Some Thessalian inscriptions (but not all, Buck 1955, p. 138) show an unusual development of the PIE interrogative pronoun *k^w*is*, namely *kis*, 'who' (Example 2.131). Other labialized velars show the expected labial or coronal outcomes in Thessalian, as in Examples 2.124-2.132. Dunnett (1970) argues that paradigmatic generalization of a velar onset, itself the result of convergence of *k^w and *k before consonantal *y within the interrogative pronoun paradigm, resulted in the Thessalian form. The argument rests on the addition of grammatical endings to the zero-grade form *k^w*i-*. In the nominative, this would have yielded the proto-form *k^w*is*, reflexes of which are attested elsewhere in Greek as *tís*. However, if the grammatical ending was vocalic, the *i would become consonantal and the onset would be a cluster of a labialized velar and a palatal; for example, the neuter plural ending was *-*a*, so the neuter plural proto-form of the interrogative pronoun would be *k^w*ya*. Clusters of a labialized velar followed by *y seem to have followed the same developments as plain velars followed by *y (Sihler, 1995, p. 190ff.); this account explains the Boeot. form *tá* and Megarian *sá* 'why, how'. Hom. *ássa* and Att. *átta* [acc.-nom. neut. pl.] 'who' (by reanalysis of the common phrase *hoppoiá ssa* 'of what sort', see e.g. Beekes 2010) are then analogous to other words in which earlier *ky developed to /ss/, or /tt/ in Attic-Ionic, intervocalically. In Thessalian, the change of *k^w to *k in those parts of the paradigm with consonantal *y may then have been generalized to the rest, giving /k/ throughout the paradigm, whilst other dialects generalized /$\mathrm{k^w}$/. The argument suffers from the need to explain why the new /ky/ cluster did not subsequently develop to /s/ as would be expected; Dunnett (1970) proposes that these expected forms, which no longer matched the rest of the paradigm due to their non-velar onset, were replaced by alternative forms with an /n/ which are attested in the alphabetic period: e.g. *kines* [nom. pl. m./f.]. An alternative explanation for Thess. *kis* is pressure from the negativized variant in early Greek, *ouk*^{w}*is* 'nobody' (later *oútis* in other Gk. dialects), in which *k^w may have undergone dissimilation (Lejeune, 1972, §31) under the influence of a neighbouring *u (see §2.6). This seems a possible but perhaps not very likely explanation; the continued productivity of **$k^w$is* in early Greek was apparently enough to prevent **ouk^{w}is* undergoing 'boukolos' type dissimilation at all in other dialects, let alone spread a dissimilated variant through the whole interrogative pronominal paradigm. Regardless of the true explanation, it is clear that the development of a labialized velar to a non-labialized velar in this position was not a systematic phonetic change in Thessalian.

2.5.7 West Greek

The West Greek grouping serves to separate those dialects traditionally referred to as 'Doric' into a group still referred to as Doric (Saronic, Argolic, Laconin/Messenian, Insular, Crete) and a subgroup of the Doric dialects now termed North-West Greek (Phocian, Locrian, Achaean, Elian) (Colvin, 2010). West Greek shows developments conditioned by the following vowel, as described in Table 1.1.

2.5.8 Strange alternations and the 'substrate' hypothesis

Some cross-dialectal alternations between coronals, velars, and labials do not have straightforward explanations as reflexes of Indo-European labialized velars, but do seem to descend from common proto-forms with labiality and velarity.

(2.136) Att.-Ion. *gép^{h}ūra*, Boeot. *bép^{h}ura*, Cret. *dép^{h}ura*, Lacon. *díp^{h}oura* 'bridge'. (Also *blép^{h}uran* attributed to Thebans; Beekes 2002, p. 16–17).

(2.137) Att.-Ion. *auk^{h}ḗn*, Hesych. *amp^{h}ḗn*, Aeol. *ámp^{h}ena* [acc.] (Theocritus, 30.28), *aúp^{h}ēn* (Jo. Gramm. *Comp.* 3, 16), 'neck, throat'.

(2.138) *daúkos*, type of plant; also Hesych. *dauk^{h}mós*, a gloss for *daúkos*; *dauk^{h}mós* 'Cretensis'; *dáp^{h}nē* 'laurel'; Thess. and Cyp. *daúk^{h}n-* 'laurel' (in compounds).

Based on the attested forms from Boeotian and Cretan, Example 2.136 looks like it could originally have had a labialized velar as the first obstruent, given the alternation between Boeot. /b/ and Cret. /d/, which are the expected outcomes of a labialized velar before **e* in these dialects. The /g/ in Attic indicates velarity, but is difficult to explain since the normal outcome of $*g^w$ before **e* in Attic-Ionic was /d/. There is a cognate in Arm. *kamurǰ*, although Frisk (1972, p. 303) points out that the Arm. /m/ is not the expected outcome of an inherited $*b^h$ (required to fit the Greek form) and Chantraine (1980) believes the word is not Indo-European. Beekes (2002) argues that the original meaning of the word in Greek was 'beam', and that both the Greek and Armenian words are connected to the (non-Indo-European) Hattic *hammuruwa* 'beam', either via borrowing from Hattic or via borrowing from a pre-Greek language into all three.

Examples 2.137 and 2.138 show reasonably similar alternations between an obstruent, and a combination of an obstruent preceded by /u/. A form with a rounded vowel followed by a velar, *aukʰḗn*, alternates with a semantically identical form with a nasal and a bilabial, *ampʰḗn* 'neck'; both may descend from an PIE form **h₂emǵʰ-u-* 'narrowness', cf. Skt. *aṃhú-* 'narrow' and Go. *aggwus* 'narrow'. There is also a possible link to Armenian *awji-k* 'neck', but the form is problematic as it does not show the expected Arm. /z/ from $*\acute{g}^h$, nor does it reflect the usual merger of labialized velars, velars, and palatals after **w* in Armenian; nor is the loss of the nasal accounted for (Clackson, 1994, p. 107ff.). Martirosyan (2008) proposes a developmental path in which a zero-grade root $*h_2ng^{wh}$- gives Proto-Armenian $*aug^h$- and then Arm. *awj-* through regular sound change and coarticulatory labialization of the vowel. Greek *ampʰḗn* can then descend from the same zero-grade form, with assimilation of the nasal to the following labial. Of course, this does not help with the Greek forms with /u/; another zero-grade form, adapted from a substrate language and realised as $*h_2u\acute{g}^h$, could lead to *aukʰḗn*, with Arm *viz* 'neck' derived from an equivalent full-grade $*w\bar{e}\acute{g}^h$-. Clearly there is no satisfactory explanation which can succinctly account for all of the Greek forms or connect them indisputably to Armenian.

The words for 'laurel' provided in Example 2.138 must be related to one another; there are no obvious Indo-European cognates. The phonological problem is strikingly similar to that of the 'neck' words in Example 2.137. Next to the forms provided in Example 2.138 is the following gloss from Hesychius:

λάφνη· δάφνη, Περγαῖοι

láp^hnē: dáp^hnē, Pergaíoi

Alternations of /l/ with /d/ are also indicated by Mycenaean Greek, e.g. <*da-pu$_2$-ri-to-jo*> for *laburint^hoio* (Chantraine, 1980); the /l/ connects the other Greek forms to Lat. *laurus* 'laurel'. The suffix *-nt^h-* is generally regarded as indicative of a borrowing from a pre-Greek substrate language; its presence in this word, in combination with the /l/-/d/ alternation which is also a feature of borrowed words, lends further support to a borrowing hypothesis. The presence of labialized stops in a pre-Greek substrate language is suspected for a number of reasons. First, the apparent existence of a series of syllabograms in Linear B representing obstruents with concurrent labialization (the <qV> series, as well as signs for <*dwe, dwo, nwa, swa*> etc.; PIE is not thought to have had phonemically labialized resonants) suggests that the undeciphered non-Greek language which Linear A transcribes had a meaningful opposition between labialized and non-labialized obstruents. Second, there are other examples of an obstruent alternating dialectally with an obstruent preceded by /u/, similar to those described in Examples 2.136-2.138.

(2.139) Att. *aúlaks*, Dor. *ṓlaks*, Lacon. *euláká̄*, Hesych. *aūlák^ha* 'plough'.

(2.140) *aurosk^hás* 'vine', Hesych. *arask^hádes* 'last year's vines'.

The words in Examples 2.139 and 2.140 have all proven impossible to explain as Indo-European forms (Beekes, 2010; Kuiper, 1968). Other words show a similar alternation involving an unexplained /u/, but in positions adjacent to obstruents

(Examples 2.141-2.142); in the case of Example 2.143, the alternation is with /k/, /p/, and a round vowel /o/.

(2.141) *tʰálpō* 'heat, warm', *tʰalukrós* 'hot', Hesych. *tʰalúpsa* 'heat'; other derivatives are listed in Hesych.

(2.142) *Puanépsia*, *Puanópsia*, *Panópsia*, *Kuanópsia*, an Athenian festival (Kuiper, 1968, p. 274).

(2.143) Att. *párnops*, Aeol. *pórnops*, elsewhere *kórnops*; Hesych. *kornṓpides* for *kṓnōpes*, also Hesych. *akornoí* and *okornoí* in Aristophanes.

An alternative explanation is that all these forms descend from a non-Indo-European word borrowed from a substrate language at a time which precluded them from undergoing the expected sound changes within the phonological chronology that inherited words experienced. It is perhaps not unexpected that the labial outcomes match the reflexes of the inherited labialized velars, since sound changes affecting labialized velars tend to produce bilabials in numerous languages (Chapter 5). It may be the case that these words were borrowed into Greek at a time when the inherited labialized velars had diverged in some dialects but not others, and were subsequently generalized across linguistic areas; or at a time when labialized velars had undergone some of their conditioned developments but not others (there are apparently no coronal outcomes in these forms). A possible complication to this analysis is the word *ántʰrōpos*, which was without a realistic PIE etymology for a long time. Suggestions are summarized in Beekes (1995, pp. 13–15), with the most plausible proposing some combination of **h₂ner-* 'man' and **h₃ēkʷ-s-* 'face', but none are either phonologically or semantically satisfactory (Kuiper, 1968, p. 275–276). The attestation in Mycenaean Greek of <*a-to-ro-qo*> (PY Ta 722.1), *antʰrōkʷōi* [dat. sg.] 'man', confirmed the labialized velar but did not add any evidence in favour of the PIE etymology. Hesychius' gloss *drṓps* for *ántʰrōpos* is plausibly related to *ántʰrōpos*, if the analogy is made to other non-IE Greek dialect words in which /tʰ/ alternates with /d/ (e.g. Hesych. *dálagkʰan* for *tʰálassan* 'sea', Kuiper 1968). Similarly, other potentially borrowed words have an occasional epenthetic vowel that seems to appear at random, as in Example 2.141, or an epenthetic nasal as in Examples 2.137 and 2.138. Thus a range of shared alternations, inexplicable in Indo-European terms, indicate borrowings of words which were articulated with labialized velars in their own languages (or at least sounds which were analyzed by Greek speakers as labialized velars). This does not provide much help with the phonological development of the Greek forms, though.

A chronological problem is posed by Myc. <*qe-to*> (REF.), *k^w^etos*, cf. later Gk. *píthos* 'jar'. The word shows alternations of a similar type to those outlined previously in this section, such as inconsistency in aspiration (e.g. *p^h^id-* (Poll.), Dor. *pisánka*) and alternation of /d/ with /t^h/ (e.g. *p^h^idáknē* in some authors). The word was borrowed into Mycenaean Greek at a time when the inherited labialized velars had not yet completed their divergent developments and is represented by the same syllabogram as the inherited PIE sounds, but in later Greek it does not show the coronal outcome that would be expected after a front vowel. It may be that the developments of each borrowed word depended upon the chronology of its borrowing over a long period of bilingualism or contact of differing degrees of closeness in different areas, resulting in variable outcomes across Greek-speaking communities. Alternatively, productive bilingualism may have kept the borrowed labialized velars separate in some way from their inherited counterparts, or meant that they were realized phonetically in a manner that was different but similar enough to use the <qV> series to transcribe them in Linear B.

2.5.9 Dissimilation near labials

There are some reconstructed PIE roots that seem to show velar developments for both $*k^w$ and $*kw$ in positions where one would expect to see a labial outcome in Greek. In example 2.144, the justification for reconstructing the sequence $*kw$ rather than a velar stop $*k$ appears to rest solely on Prussian *po-quelb-* 'kneeling'; this connection seems semantically difficult, and there is no evidence of labiality in the other proposed cognates, making it something of a stretch to draw any conclusion about PIE labialized velars from this root. More problematic are examples 2.145 and 2.146, in which PIE labialized velars show a velar reflex in Greek. Pokorny (1959) and Beekes (2010) suggest that dissimilation due to similar phonetic environments resulted in delabialization of the labialized velar and explain these exceptional developments. However, the phonetic environments are not exactly parallel; Gk. *karpós* must come from the zero grade, meaning the labialized velar was preceding an unstressed syllabic resonant followed by a labial; for Gk. *kólpos*, presumably from o-grade **k^w olp-*, the labialized velar was followed by a rounded vowel, a consonantal resonant, and a labial. On the other hand, it is impossible to know the precise realizations of the resonants in relation to these different environments, and there is clear evidence of similar distance dissimilations in other languages (e.g. Ital. *cinque* from Lat. *quinque*, see e.g. Elcock 1975, p. 73).

(2.144) PIE **kwelp-* 'stumble' gives Gk. *kálp-* 'trot' (from the zero grade), cf. OHG *holp-* 'jump', Lith. *klump-* 'stumble', and possibly Prus. *po-quelb-* 'kneeling' (Frisk, 1972) although Chantraine (1980) does not make the connection and Beekes (2010) argues against it as it leaves the Gk. vowel quality unexplained.

(2.145) PIE **k^w erp-*, **k^w rp-* 'turn' gives Gk. *karpós* 'wrist', cf. Goth. ƕ*aírban* (Lejeune 1972, p. 72, n. 3; Chantraine 1980; Frisk 1972).

(2.146) PIE **k^w elp-*, *k^w olp-* 'curve' gives Gk. *kólpos* 'breast', cf. OHG *walm* and OIsl. *huelf-* (Lejeune 1972, p. 72, n. 3; Chantraine 1980).

2.6 The 'boukolos' phenomenon

2.6.1 Introduction

In the vicinity of a rounded vowel, labialized velar stops occasionally show historical delabialization even when reflexes of labialized velars elsewhere in the language are labial. This effect is observable not only in historical sound change, but in synchronic phonological alternations induced by morphological operations. The name ‘boukolos’ originates in the field of Indo-European historical linguistics, but labialized velar dissimilation near rounded vowels is not limited to the Indo-European languages, as will be discussed further in Chapter 5.

Greek and the origin of ‘boukolos’ as a descriptive term

The quality of the Greek /u/ vowel is discussed in §2.4.2, in which it was concluded that the vowel was a high back rounded vowel similar to [u]. The Indo-European vowel was most likely similar to the Greek vowel.

There are several proposed examples of ‘boukolos’ style sound changes, but some are more secure than others. In Greek the dissimilation appears to have applied universally, with the name ‘boukolos rule’ arising from the most secure example of this type of change (see Example 2.147).

(2.147) PIE *g^w*ou-*k^w*ol-* gives Gk. *boukólos* ‘cowherd’, first attested in Myc. <*qo-u-ko-ro*> for g^w*oukolos*, from a compound of *g^w*ou-* ‘cow’ and *k^w*ol-* ‘turn’.

The example is noteworthy in that examples of other compounds formed from the PIE root *k^w*el-* ‘turn’ retain the labialized velar and show labial outcomes, as would be expected (Examples 2.148-2.149).

(2.148) Gk. *aipólos* ‘goatherd’ from PIE **aiǵ-* ‘goat’ and *k^w*el-* ‘turn’.

(2.149) Gk. *hippopólos* ‘herding horses’ from PIE **eḱw-* ‘horse’ and *k^w*el-* ‘turn’

Additional examples of ‘boukolos’ dissimilation within Greek are presented in Examples 2.150-2.152.

(2.150) PIE *$leng^{wh}$- 'light', appears in Gk. *elak^{h}ús* 'small' with velar development under the influence of the suffix *-u-* (cf. Skt. *laghú-* 'light', Av. *ragu-* 'quick'); the same root is in Gk. *elap^{h}rós* 'light' with the suffix *-*r*-, in which the labialized velar shows the expected labial development.

(2.151) PIE *h_1ewg^{wh}- 'speak solemnly', gives e.g. Gk. *eúk^{h}omai* 'pray', first attested in Myc. *e-u-ke-to*.

(2.152) PIE *b^hlewg^w- 'blow, swell' gives Gk. *p^{h}lúks-*, cf. Lat. *flūct-* (Pokorny 1959; disputed in Weiss 1994)

Mycenaean Greek

Notably, juxtaposed /u/ and /kw/ in some Mycenaean compounds do not show this dissimilation (Examples 2.153-2.159). The absence of dissimilation could be due to the compounds post-dating a universally applied sound change Palmer (1980) or to analogical retention of the labialized velar, due to the continued use of the individual lexemes that made up the compound. Alternatively, pronunciation of /kw/ adjacent to /u/ could in fact have been similar or even identical to /k/ adjacent to /u/, but the continued existence of constituent words with /kw/ led to phonological analysis of the word with /kw/, which is reflected in the orthography. For example, in the case of Example 2.153, phonological analysis of *ouk^{w}e* 'and not' as containing a labialized velar may have continued, despite phonetic ambiguousness of the distinction between /ukw/ and /uk/, due to the existence of the early Greek particle *k^{w}e* 'and'. Examples 2.154-2.159 all show compounds which have parallels in which the labialized-velar-containing constituent occurs adjacent to segments other than /u/, providing opportunities for phonological parsing of labialization as a feature of the stop as well as the vowel in these examples. There is a precedent for such an argument; Hale (2012) presents a detailed study of Marshallese sound changes in which the phonological identity of a vowel alters whilst its phonetic realization remains constant.

(2.153) Myc. *o-u-qe* for *ouk^{w}e* 'and not', cf. later Gk. *oúte*.

(2.154) Myc. *qo-u-qo-ta* (KN L 480b) for $g^w oug^w ot\bar{a}i$, personal name; cf. later *boubotas* 'giving pasture' (Pindar *Nem.* 4.52).

(2.155) Myc. *po-ru-qo-ta* (PY Jn 845.2), *Poluk*^{wh}*óntas*, personal name; cf. Hom. Polyphontes (Aura Jorro, 1993).

(2.156) Myc. *po-ru-qo-to* (PY An 128.2), perhaps *Polúg*^{w}*otos*, personal name; cf. later *Polúbotos*, *Poluboútēs* (Aura Jorro, 1993).

(2.157) Myc. *su-qo-ta* (PY Ea 822 etc.), *sug*^{w}*otas* 'pig farmer'; cf. later *subòtas* 'pig farmer' (Aura Jorro, 1993, p. 304–305).

(2.158) Myc. *e-u-ru-qo-ta* (KN V 147.2), *Eurug*^{w}*otas*, personal name (Aura Jorro, 1993, p. 265).

(2.159) Myc. *a-tu-qo-ta* (KN B 799.8), personal name on a list of personal names, but the exact transcription is not certain; possibly *Atuk*^{wh}*ontās*, cf. later Gk. name *Átus*, or *Artug*^{wh}*óntas* (Aura Jorro, 1993, p. 122).

2.6.2 Other Indo-European languages

Additionally, there is some evidence of 'boukolos'-type dissimilations in other IE languages; see examples 2.160-2.163. The question of whether or not 'boukolos' dissimilation was widespread in PIE is difficult to answer. One important consideration is that if labialized velar dissimilation near **u* did occur early in PIE, the evidence for it would naturally have been eliminated in all languages, as only plain velars would continue to exist in proximity to **u*. Weiss (1994) argues that this is indeed the case and that the claim is supported by a unexpectedly high proportion of PIE roots that feature a plain velar in proximity to **u*, presumably to be interpreted as the descendants of delabialized labialized velars.

(2.160) OIr. *bóchaill* 'cattle dog' and Wel. *bugail* 'herdsman', from PIE *$g^w ou\text{-}k^w ol\text{-}$ 'cowherd' (Vendryes, 1959).

(2.161) Gk. *hugiḗs* 'healthy'; either a compound of *$h_1 su$- 'well' and *$g^w ih_3$- 'live', with potential cognate of Av. *hu-ǰyā-ti-* 'good life' (Frisk, 1972); alternatively, a compound of *h_2*iu*- 'life, eternity' and *g^w*ih*$_3$- 'live', with cognates Av. *yauuaē̆ǰī-* 'living forever', Lat. *iūgis* 'everflowing', Goth. *ajukdūþs* 'eternity' (Weiss, 1994).

(2.162) PIE *b^h*ouk*^{w}- 'flying insect' giving Lat. *fūc-* 'drone'; OE *béaw* 'gadfly' with labiality retained (Pokorny, 1959).

(2.163) PIE *b^h*leug*^{w}- 'flow' > Lat. *fluxi*, Gk. *p*^{h}*lúzdō* 'seethe'; however Rix (2001) and Derksen (2008) identify an alternate root *b^h*leuH-* giving OCS *bljuj-* 'vomit' and Gk. *p*^{h}*léō* 'flow over'.

Conversely, there are some apparent examples of PIE roots which have been reconstructed with a *u vowel adjacent to a labialized velar for which reflexes in the daughter languages do not demonstrate velar developments (examples 2.164-2.165) indicating that dissimilation did not take place.

(2.164) **e-wewk*^{w}*-om* reduplicated aorist of 'speak', Gk. reflex *eípon* 'spoke'.

(2.165) **luk*^{w}- 'wolf', Lat. *lupus* 'wolf', *uulpes* 'fox'.

Weiss (1994) rejects examples 2.164 -2.165, arguing that the 'boukolos' changes took place in PIE, rather than in Greek as was traditionally supposed. His analysis is credible on many points, particularly the principle argument proposing the derivation outlined in example 2.161. The reconstruction of a compound *h_2*iu-g*^{w}*ih*$_3$*s*, giving later PIE *h_2*iu-gih*$_3$*s* via labial dissimilation, is more believable on phonological, semantic, and comparative grounds than earlier suggestions, which required a phonologically anomalous development of *h_1*su-*. These earlier reconstructions contradicted the reasonably reliable reconstruction offered in Example 2.166 and the conflicting evidence provided by its development in numerous other compounds such as that described in Example 2.167.

(2.166) Gk. *eu-* 'well' from PIE *h_1*su-* 'well, good', cf. Skt. *su-* 'good' and Hitt. *āššu-* 'good'.

(2.167) Gk. *eúbat-* 'accessible', from PIE *h_1*su-* 'well' and *g^w*m-* 'go'.

Thus, the derivation of the Greek, Avestan, Latin, and Gothic forms from a relatively late PIE *h_2*iu-gih*$_3$*s*, in which *g^w had already undergone a loss of contrastive labialization due to being adjacent to **u* in the formation of the compound. However, some of the reconstructions proposed as proof that early PIE labialized velars underwent ‘boukolos’ type delabialization are less convincing than *h_2*iu-g*^{w}*ih*$_3$*s*. Pokorny (1959) reconstructs a root that has **u* adjacent to a labialized velar, b^h*ouk*^{w}- ‘flying insect’ (Example 2.162) to explain Latin *fūcus*, suggesting it is a cognate of OE *bēaw* ‘gadfly’; the labialized velar would then explain the velar stop of the Latin form and the labialization in the OE form. Pokorny’s reconstruction is challenged by an alternative derivation of Lat. *fūc-* from **bhoik-* (Weiss, 1994). This would make Lat. *fūc-* cognate to OCS *bičela* ‘bee’ from earlier *b^h*ik-el-* and OIr. *bech* ‘bee’ [m.] from PIE *b^h*ek-* (de Vaan, 2008). These derivations would have to come from *e*-grade or *i*-grade forms of the root, whilst Lat. *fūcus* would then be from an ablaut form with **oi*; that form remains without any direct cognates with the same ablaut formation. Thus, though the words are probably related in some way, their etymology and relationship to one another, as well as the existence or not of a form *b^h*ouk*^{w}-, remains uncertain.

Another possible example of ‘boukolos’ type delabialization in PIE is Lat. *ūuidus*, which is reinterpreted by Weiss as a reflex of the zero-grade of the root **ug*^{w}- ‘wet’ in which the labialized velar was lost through ‘boukolos’ dissimilation and subsequently restored by analogy to the full grade. However, as Weiss acknowledges, there are numerous possible derivations for *ūuidum*: a second possibility is that *ūuidus* came from a different PIE root without a labialized velar, with the addition of Latin adjectival ending *-idus* giving *ūdus* by regular sound change, then *ūuidus* as a hyperarchaic back-formation. On the other hand, the order of attestation suggests that *ūdus* is a later contraction of *ūuidus*, and whilst the explanation works phonologically, *-idus* was a productive Caland suffix, and there is no evidence for **weh* to have a Caland system whereas there is plenty for **weg*^{w}-. A third possibility is that an early unattested noun *ūmos* resulted from the zero grade of the **weh*$_1$

root, producing *ūmor* analogically, then *ūuidus*, by some more analogy, presumably to *ūmidus*. Neither of these latter explanations seem any more convincing than a straightforward account of an early form **ug^w-ido-* giving *ūvidus*; the only benefit of introducing multiple reanalyses or archaic back formations is to support the argument that the 'boukolos' change was universal in PIE. Similarly, Lat. *lupus* 'wolf' as a reflex of **luk^w-os* is rejected as an example of labialized velar retention on the basis that Latin was unlikely to have borrowed such a basic lexeme from a language of the Osco-Umbrian family—in which $*k^w > p$, e.g. Umbrian *petur-* 'four-' from **k^{w}etwor-* 'four'—due to the equally likely preponderance of wolves in Latin-speaking regions (Weiss, 1994, p. 140). However, there is good evidence that Latin borrowed equally basic vocabulary from other Italic languages: directly comparable is *bōs* 'cow' and *rūfus* 'red', both borrowed from a Sabellic language de Vaan (2008). Therefore it seems entirely plausible that *lupus* could have been borrowed. An alternative explanation is that *lupus* comes rather from a root **wlp-, lup-* (cf. Av. *urupis* 'dog', Arm. *aluēs* 'fox', Gk. *alōpos* 'fox'); but there already exists a reflex of this root in Latin (*uulpes* 'fox') so the explanation requires borrowing, which Weiss has already argued is implausible for this word (there are numerous examples in the world's languages in which communities have borrowed terms for everyday objects or concepts from dialects or languages with which they had contact; English pronouns beginning *th-* were borrowed from Norse, for example).

The etymologies of Lat. *fūcus* and *ūuidus* are too uncertain to make them reliable evidence for 'boukolos' type changes having been regular in early PIE, and attempts to prove otherwise involve too much conjecture to be overwhelmingly convincing. On the other hand, Weiss's etymology for **h$_2$iu-g^{w}ih$_3$s* is much more convincing than earlier suggestions and is well supported by comparative evidence from daughter languages. Although this reconstruction alone does not provide conclusive evidence that all instances of **u* adjacent to labialized velars underwent delabialization in PIE, it suggests that it did in at least one case, and reliable contradictory examples in which labialized velars persist adjacent to *u are rare.

2.6.3 'Boukolos' dissimilation in non-IE language families

There are numerous examples of similar dissimilation taking place in language families outside Indo-European. Examples are provided in §5.4.

2.7 Greek outcomes of PIE clusters with **w*

In word-initial position, **tw* clusters inherited from PIE became sibilants in all dialects except Doric, in which they became coronals (Examples from Sihler 1995, p. 186):

(2.168) PIE **twe* 'you', gives Dor. *té* 'you' (perhaps also /twe/, see the Hesychian gloss τρέ), elsewhere *sé* 'you'.

(2.169) PIE **twei-* 'excite' gives Gk. *seíō* 'shake'

Word-internally, outcomes for **tw*, **ky*, and *$k^w y$ generally coincide, giving *-ss-* in most dialects and *-tt-* in Attic and Boeotian (Sihler, 1995, p. 186) as in Example 2.124. However, there are some exceptions such as Att. *episeíō* 'shake at' (see Example 2.169) and Lesb. *pésures*, Dor. *tétores* 'four' from PIE *$k^w etwor$-. After an aspirated coronal *t^h, **w* is lost, e.g. $ort^h ós$ and Dor. $bort^h o$- 'straight' from PIE *$wort^h wos$.

Clusters of a resonant followed by **w* are retained in Mycenaean (see §2.5.2 for examples) and in many early dialect inscriptions. The /w/ was subsequently lost in all dialects except Tsakonian, often with compensatory lengthening of the preceding vowel (Examples 2.170-2.171).

(2.170) Inscriptional *ksenwos* 'stranger', cf. later Att. *ksénos*, Hom. *kseínos*, Dor. *ksēnos* 'stranger'.

(2.171) Inscriptional *korwa* 'girl', cf. later Att. *kórē*, Hom. *koúrē*, Dor. *kōra* 'girl'.

2.8 Divergent outcomes dependent upon voicing and aspiration

The material presented in this chapter so far can be condensed into three main observations concerning Greek outcomes: first, the regular outcome of $*k^w$ before $*i$ is coronal in all the dialects. Second, the outcomes of $*k^w$, $*g^w$, and $*k^{wh}$ are coronal before $*e$ in all the dialects except Aeolic. Third, the outcomes of $*g^w$ and $*k^{wh}$ are labial before $*i$. However, the details on some of these points are problematic. There is no obvious reason, whether through conditioning environments, structural constraints, or analogical considerations why the voiced and voiceless aspirated labialized velars do not both become coronals before $*i$. This has led to several varying analyses. For example, Meillet (1894) and Sheets (1975) argue that all the coronal outcomes of $*k^w$ before $*i$ are analogical, and that $*k^w i$ to /pi/ was in fact the regular development; Meillet's argument is that the Greek /e/ became pre-palatalized /^{j}e/, triggering assimilatory palatalization of the labialized velar before reverting back to /e/. This seems possible but unlikely, and at any rate the examples presented in this chapter in favour of $*k^w i$ to /ti/ sound changes are too numerous and reliable, relative to the changes of $*k^w i$ to /pi/, to support their analysis as exceptional. The problem then remains why mid front vowels should create a palatalizating environment for voiced and voiceless aspirated labialized velars, but not high front vowels.

Analysis of the voiced and voiceless aspirated labialized velars is complicated by the cross-dialectal attestation of labial and coronal outcomes in both vocalic environments. Possible evidence for coronal outcomes of $*g^w$ is presented in Examples 2.172-2.175. The etymologies presented in Examples 2.172-2.174 are from a personal communication with Lucien van Beek.

(2.172) *aídios* 'eternal' from earlier $*aiwi\text{-}g^w ios$.

(2.173) *ídios* 'private' from earlier **swe-g^w ios.*

(2.174) *Antídios*, personal name, cf. Hom. *antíbios* 'opposing'.

(2.175) *endediōkota* 'established', Heracl., from PIE root **g^w i-* 'live'.

Examples 2.172-2.174, in van Beek's analysis, contain coronal reflexes of the PIE root **g^w i-* 'live'. Thus, coronal reflexes of voiced labialized velars would be attested, but the other examples in which there is a labial outcome remain unexplained.

For the voiceless aspirated labialized velar, there is only one example of a PIE word containing this sound appearing in Greek:

(2.176) **eg^{wh} i* 'snake', giving Gk *óp^h is* 'snake', Skt. *áhi-* 'snake', Av. *aži* 'snake'; Pokorny (1959) also connects Illyrian *abeis* (Hesych.).

Chantraine (1980) observes that the word is subject to taboo variation (cf. the word for 'wolf' dicussed in §2.6.2) and connects it to the Greek word *éŋk^h elus* 'eel'.

(2.177) Gk *éŋk^h elus* 'eel' and Lesb. *ímbēris* 'eel' (Hesych.), cf. Lat *anguis* 'snake', Pruss. *angis* 'snake'; possibly also a connection to Arm. *iž* 'snake', Lith. *ungurỹs* 'snake'.

Thus the situation is somewhat complex, and many theories have been proposed explaining the discrepancy. One explanation is simply that the 'snake' example is not a reliable indicator of the behaviour of aspirated labialized velars before /i/, and that the lack of good examples of $*g^{wh}$ in this position means no sound conclusion can be drawn about how it would or would not have developed in Greek. However, there are good examples of labial outcomes of $*g^{w}$ before $*i$ that still have to be explained. Allen (1957) approaches the problem from a structural and phonetic perspective; it is suggested that, whilst both /e/ and /i/ were sufficiently 'palatal' to invoke a phonetically motivated palatalization of the labialized velars, only /i/ is palatal enough for coarticulatory palatalization to avoid being reanalyzed as a characteristic of the stop rather than the vowel. This is an interesting argument but is complicated by the cross-linguistic tendency for obstruents to palatalize before /i/ more often than before /e/ (Stephens & Woodard, 1986). No explanation is offered for why this would be different in Greek when, presumably, the same mismatched opportunity for phonologization would exist in other languages. However, all previous studies into the phonetics of velar palatalization have been conducted on velars without contrastive secondary labialization; it is possible that some aspect of the phonetics of labialized velars may contribute to an analysis of why labialized velars might have more readily palatalized before mid front vowels than before high front vowels.

2.9 Conclusions

Stops Based on the evidence from daughter languages presented in this chapter, it is most likely that the Indo-European labialized velars were velar stops with secondary labialization, similar to the surviving labialized velars in Latin. For some sounds there are very few available attestations, making it difficult to draw sound conclusions about their developments (e.g. voiced unaspirated labialized velars in positions preceding back round vowels). There is a lack of conclusive evidence to confirm whether or not clusters of /k/ and /w/ were phonologically contrastive with /k^w/ in the *centum* languages, although divergent development in the *satem* languages indicate that the stops and clusters were distinct in an earlier stage of Indo-European.

Vowels The conditioning environments for labialized velar developments in Greek can be described as preceding front vowels, back or central vowels, back rounded vowels, and consonants. Whilst there are several competing hypotheses concerning the PIE vowel inventory, the Greek vowel inventory is more straightforward and it is likely that the vowels /a, o/ and /e, i/ were present and caused different outcomes for a preceding labialized velar. The vowel /u/ was a back rounded vowel in early

Greek and presumably also in PIE.

3

Acoustic characteristics of labialized velars

3.1 Introduction

The theoretical framework outlined in Chapter 1 emphasises the role of synchronic phonetic variation in the inception of historic phonological sound changes; it was shown that previous studies have successfully identified sound changes caused by cross-linguistic tendencies towards specific variations in production and perception and supported conclusions with acoustic and perceptual data. Within this framework I proposed to use primary phonetic evidence to answer questions relating to the inception of specific sound changes involving labialized velar stops in Indo-European and ancient Greek. It goes without saying that phonetic evidence from these languages is no longer available; therefore, the current chapter will use as evidence phonetic data from proxy languages (methodological justification and precedence for the use of proxy languages is outlined in §1.5). British English and Western Zapotec are used as proxy languages. This enables a comparison of labialized velars as phonologically complex clusters (a possible analysis in British English) and as monophonemic segments (a possible analysis in Western Zapotec); as was discussed in §2.3, the phonological status of labialized velars as stops or clusters in PIE and the daughter language is not fully understood. Labialized velars in British English are voiceless and aspirated, and in Western Zapotec they are voiceless and somewhat aspirated in most instances. The aspirated labialized velars are not well attested in Ancient Greek, and so it is difficult to establish their reflexes; in §2.8 it was observed that voiceless aspirated labialized velars may have labial outcomes before front vowels in Greek and coronal outcomes before mid front vowels, in contrast to the coronal outcomes in both positions undergone by voiceless unaspirated labialized velars. Thus, using aspirated labialized velars in the acoustic analysis will allow for comparison of their behaviour in front and mid vowel environments, informing a discussion of the possible reasons for the divergent developments in Ancient Greek.

This chapter provides an acoustic characterization of labialized velar stops, identifying their crosslinguistic attributes in two languages and the key characteristics that distinguish them from other stops. Acoustic analysis of audio material from British English and Western Zapotec will contribute to the discussion. Identifying the important acoustic cues to labialized velar identification will help to clarify potential mechanisms for perceptually motivated historical developments. Analysing the ways in which the phonetic realization of labialized velars differs in various vocalic environments, and identifying cross-linguistic similarities and differences, will allow generalizations applicable to the Ancient Greek sounds and inform potential mechanisms for the conditioning of the sound changes.

3.1.1 Acoustic measurements

Describing the acoustic cues available to speaker-listeners to distinguish labialized velar stops in different vowel environments will make it possible to identify potential motivations for perceptually conditioned sound changes. Speech sounds are often described dynamically by measuring formant transitions, whilst spectral slices may be used to describe an acoustic event statically. In identifying which types of variation occur in specific sequences it will be possible to suggest potential mechanisms for reparsing and reanalysis leading to phonetically motivated sound changes.

Formant frequencies During phonation, air is made to vibrate by rapid opening and closing of the glottis; the rate at which this happens is the fundamental frequency or f_0. As it passes through the vocal tract, air resonates at frequencies determined by the size and shape of the cavity; these resonances are referred to as formants. The amplitude of the sound wave is increased at these frequencies, appearing as a peak on a spectrogram. Different shapes of the vocal tract produce different resonances and thus different formant frequencies in the acoustic signal, which can be used to describe and characterize the acoustics of speech sounds. Speakers rely upon formant frequencies to distinguish speech sounds; generally the first to fifth formant frequencies are measured for human speech (F_1 to F_5), with the first and second being most relevant to distinguishing speech sounds (Olive et al., 1993, p. 80). Formant frequencies for the 'same' vowels tend to differ following different consonants owing to the natural overlapping of articulatory gestures, which results in a different configuration of the vocal tract at voicing onset. Listeners use differences in the formant transitions of vowels to identify preceding stop bursts; however, a cue being encoded in the acoustic signal and therefore available to listeners does not necessarily mean that listeners can or will use it to distinguish speech sounds. In perceptual psychology, a Just Noticeable Difference (JND) is defined as the least amount by which something must be changed for a difference to be perceptible. Research indicates that for speech sounds the JND of formant frequencies is mutable depending on the context within which a speech sound occurs (Mermelstein, 1978) and, in experiments, differs depending on how many formants are changed simultaneously (Hawks, 1994). For isolated synthetic vowel formants, JND's are reported as being as low as 1% (Kewley-Port & Watson, 1994), 2% (Hawks, 1994) or between 3–5% (Flanagan, 1955; Mermelstein, 1978; Nord & Sventelius, 1979) of the formant frequency. This suggests that humans are able to distinguish quite fine differences in formant frequencies; for example, for the F_2 frequency of a typical [i] vowel at 2400 Hz (Catford, 1988), the JND required for a listener to register a change in vowel quality might lie in the range 24–120 Hz.

Voice Onset Time Voice Onset Time (VOT) refers to the interval between the release of a stop constriction and the onset of phonation of a following voiced sound. For voiced stops, VOT is negative due to phonation beginning before the burst. VOT has been shown to vary depending on the place of articulation of the stop and on the quality of the following vowel, variation which may be caused by aerodynamic and physiological constraints, and by language-specific phonological constraints (Klatt, 1975; Lisker & Abramson, 1967; Rosner et al., 2000).

Burst spectra Burst spectra are spectral slices measured at the stop burst and show the relative intensity of points across the frequency range. Burst spectra are useful in characterizing stops as they provide a 'snapshot' of the acoustic signal at the moment when the articulators are closest to their target positions for a specific obstruent, but are nevertheless affected by coarticulation to adjacent sounds. Analysis generally focusses on the distribution of energy, with peaks on the spectra representing frequencies at which the amplitude is boosted by the resonances of the vocal tract; spectra may also be diffuse (energy is distributed relatively evenly across the frequency range) or compact (energy is concentrated at specific frequencies). Overall loudness or quietness of the burst can be shown by finding the root mean square of the amplitude across the frequency range.

3.2 Literature review

Due to the common analysis of labialized velars as complex clusters rather than phonologically simple stops, examples of acoustic analyses of these sounds are scarce. Many phonetic or acoustic descriptions of languages which have labialized velars in their inventories address other stop consonants at length whilst devoting little space to labialized velars (Miller & Nicely, 1955; Maddieson et al., 1993; McDonough & Ladefoged, 1993b,a; Maddieson et al., 1996). Elsewhere, discussion of sounds with both labial and velar articulation is primarily focussed on the labiovelar approximant [w] (Ladefoged, 2001). However, discussions of the separate velar and labial components are widespread, and there are relevant discussions on the phonetics and phonology of secondary articulations in general and of doubly articulated stops.

3.2.1 Velars

Formant frequencies The shape of the vocal tract during velar constriction causes similar second and third formant frequencies, resulting in the appearance of convergence on a spectrogram before and after constriction known as the 'velar pinch' (Ladefoged, 2001; Ali et al., 2001). This effect is more prominent in front vowel environments than in back vowel environments (Olive et al., 1993, p. 147, 151) suggesting that it results from a prevelar place of articulation.

Voice Onset Time VOT is typically longer for velar stops than for other single articulation stops, with the exception of uvular stops (Cho & Ladefoged, 1999, p. 219, Lisker & Abramson, 1964). The increased VOT of velar stops can be explained in part by the relatively small space behind the constriction: for a velar stop the same volume of air is compressed into a smaller space than for a coronal or bilabial stop, increasing the pressure difference across the vocal folds and potentially resulting in a longer time for the pressure to adapt to the level required for vocal fold vibration after the closure is released. Likewise, the larger body of air resting in the vocal cavity in front of the constriction impedes the passage of air outwards when the constriction is released (Cho & Ladefoged, 1999, pp. 209–214). The relative softness of the tongue dorsum and the velum result in a greater area of contact between articulators during constriction than for other stops, which takes longer to be released (Cho & Ladefoged, 1999, p. 211). Relative inertia of the articulators may also affect VOT for similar reasons; the tongue tip, for example, has been shown to move faster than the tongue dorsum, creating the opportunity for air pressure to be reduced at a faster rate upon release of the closure (Kuehn & Moll, 1976). However, this does not seem to be the most crucial determiner of VOT as it predicts a relatively short VOT for apical alveolar stops, which is not generally the case (Cho & Ladefoged, 1999, p. 211). The duration of aspiration also mitigates these effects as causes of increased VOT, due to the relative lateness of vocal fold adduction in heavily aspirated stops (Cho & Ladefoged, 1999, p. 214); it is not clear at exactly which VOT these causes cease to be relevant, and for languages which do not phonemically distinguish aspiration there is a greater range of mean VOT values including VOTs which, in contrasting languages, would signify unaspiration. For aspirated stops, more likely causes of increased VOT are identified by Cho & Ladefoged as: 1) timing of the glottal opening, which precedes the release of aspirated stops to allow for aspiration and must then be reduced rapidly to permit vocal fold vibration; and 2) temporal adjustment which increases aspiration when the closure interval is decreased (as it tends to be for velar stops relative to e.g. bilabial stops) in order to maintain a constant duration of vocal fold opening (Cho & Ladefoged, 1999, p. 212–213).

Burst spectra Burst spectra of velars are compact, characterized by the presence of a prominent spectral peak (Blumstein & Stevens, 1979, p. 1006). Energy is concentrated in lower frequency ranges for backed [k] than for fronted [k] (Zue, 1976).

Context-specific effects Velars are subject to coarticulatory palatalization before front vowels in many languages. Ladefoged & Maddieson (1996) found that in a front vowel environment tongue placement was as much as 8 mm further forward than for a back vowel environment for Ewe velars. Similar results were found for English (Houde, 1968; Löfqvist & Gracco, 1994), for a Swedish speaker (Löfqvist & Gracco, 1994), and for German speakers (Mooshammer, 1992). Zue (1976, pp. 110–123) found that before back rounded vowels English [k] and [g] bursts had identical mean burst frequencies of 1250 Hz, whilst before front vowels the means were 2720 Hz in both stops. Back [k] and [g] were characterized by a second, lower energy peak at a higher frequency. Acoustic effects of single articulation velar fronting in front vowel environments in Czech, Hungarian, English, and Russian have been described by Keating & Lahiri (1993). Keating & Lahiri (1993, p. 89) suggest that the location of tongue contact for velars varies according to frontness of adjacent vowels: "the more front the vowel, the more front the velar". In front vowel contexts, the velar burst has its strongest spectral peak at a higher frequency than the following vowel's F_2, but not higher than the vowel's F_4 (Keating & Lahiri, 1993). In non-fronted velars, the strongest peak is found at a frequency at or below the vowel's F_2 (Keating & Lahiri, 1993). Thus front vowels are shown to increase the frequency of the highest spectral peak of the velar burst, relative to the effects of other vowels. The increase in F_2 that results from this type of palatalization has the effect of exaggerating the contrast of the following high vowel with other, less high vowels (Flemming, 2002, p. 101). Keating & Lahiri (1993) describe the relationship between spectral peaks and vocalic formants as orderly with relation to tongue fronting, i.e. a palatalized velar has a higher frequency spectral peak than a fronted velar due to the progressive reduction in size of the front cavity. Coarticulatory velar fronting has provided phonetic motivation for widely-attested historical developments of velars into affricates (Guion, 1996, 1998; Wilson, 2006).

3.2.2 Lip rounding and labialization

Formant frequencies The primary acoustic effect of lip rounding is lowering of the first three formant frequencies of a following vowel caused by protrusion of the lips and subsequent lengthening of the vocal cavity (Ladefoged 2001, p. 160 ff.; Ladefoged & Maddieson 1996, p. 358). Ladefoged & Maddieson (1996, p. 356-60) describe the acoustic characteristics of secondary labialization (but do not compare it to stop + labial clusters) concluding that the effects of labialization are concentrated more fully on the release phase of the primary constriction (cf. many languages which restrict the quality of the vowel following a labialized stop, but not the vowel preceding it) and causes lowering of the formant frequencies of the F_1 and F_2 transitions.

VOT Labial stops typically have shorter voice onset times than velar or coronal stops (Lisker & Abramson, 1967). Lisker & Abramson (1964) show that voice onset times for labial stops are shorter than for coronal or velar stops in English, Dutch, Spanish, and Hungarian. In languages with aspiration contrast, /p/ had a VOT less than /k/ and /t/ in Cantonese, Hindi, and Eastern Armenian; however, for all these languages /t^h/ had a shorter VOT than /p^h/ and /k^h/. In Marathi, both aspirated and unaspirated coronals had a VOT shorter than labials and velars.

Burst spectra Labial stop bursts are characterized by diffuse spectra with either consistent energy distribution across the frequency spectrum or a lessening of energy as frequency increases (Blumstein & Stevens, 1979). Phonologically contrastive secondary labiovelarization—unlike contextually conditioned labialization that serves to help differentiate a primary contrast, e.g. lip rounding in English /ʃ/

to increase formant frequency lowering and distinguish from /s/ with neutral lips (Mann & Repp, 1980)—is generally accompanied by tongue dorsum raising which starts slightly earlier than lip rounding (Ladefoged & Maddieson, 1996, p. 358). Spectrographic effects of secondary labialization on velars include a lower frequency spectral peak than for non-labialized velars (Suh, 2008; Denzer-King, 2013; Zue, 1976). Suh (2007) found that the effects of secondary labialization on stop burst spectra, namely lowering of the overall frequency, were greater for /k/ than for other stops in Spanish and Korean. Allen (1987, p. 113) show that burst spectra for velar stops with coarticulatory labialization (resulting from anticipatory coarticulation to a following rounded vowel) have a prominent low frequency peak that does not appear in the burst spectra of velars before front vowels.

3.2.3 Secondary articulation, double articulation, and clusters

Canonically, sounds with secondary articulations consist of a primary constriction articulated simultaneously with a secondary articulation, with the secondary constriction more weakly constricted than the primary; examples include stops with secondary labialization such as in [kʷ], secondary palatalization such as in [pʲ], secondary pharyngealization such as in [dˤ], and secondary velarization such as in [ɫ]. Sounds with double articulation consist of two simultaneous articulations with the same degree of stricture, e.g. [g͡b], [w] etc. Catford (1977b, p. 189; p. 195) recognises the difficulty of separating coordinate from secondary articulation in phonetic terms, and argues that doubly articulated stops and stops with primary and secondary articulations can be described as ‘stronger’ or ‘weaker’ degrees of the same phenomenon. This suggests that it could be useful to recognise degrees of constriction; for example, stops with phonemic secondary labialization and doubly articulated stops would be points on a gradient of closeness of constriction. Catford compares gradient palatalization, offering as an example the Cypriot Greek [pcos], corresponding to Athenian [pjos] ‘who’ (p. 195). Within this theoretical framework, sound changes of labialized velars to labials might be represented as gradual closure of the secondary constriction along this gradient, presumably via intermediate stages such as [k͡p].

Connell (1994) investigated characteristics of /k͡p/ and g͡b in Igbo, Obala, and Ibibio, and found that labial-velars had a greater lowering effect on the F_2 transition of a following vowel than plain labials; thus the labial-velars were characterized by a steeper F_2 transition (p. 469). This is illustrated by spectrograms and schematic formant diagrams in Ladefoged (1964, p. 12), in which the steeply rising F_2 transition of front vowels following a labial-velar stop can be clearly observed. In addition to this, labial-velars had a relatively short voice onset time, possibly the result of pharyngeal expansion during the velar articulation. Ladefoged (1964) observes that labial-velars can be produced with a pulmonic egressive airstream (the ‘normal’ mechanism for most languages) but may also be produced with simultaneously ingressive and egressive airstreams; differences are language-specific.

Duration and timing Double articulation stops—i.e. those in which near-simultaneous and equal degrees of constriction (in this case, complete closure) are applied to two articulators—have been shown to have closure durations not significantly exceeding those of equivalent single articulation stops in at least one language, but to have slightly longer durations in another (Ladefoged & Maddieson, 1996, p.333–334). However, in both of these languages clusters of single articulation stops with the same places of articulation, e.g. /kw/, had durations from 1.5 times to double that of the double articulation stops. DiCanio et al. (2013, p. 2239) found that Mixtec /kw/ was equivalent in duration to English /k/, rather than English /kw/, but did not compare labialized velars and clusters (due to phonotactic constraints on consonant clusters in Mixtec, Macaulay 1996, p. 26). Catford (1977b) identifies syllable or word boundaries as the determiner of increased length in stop clusters (Haggard, 1973). Connell (1994) found that Igbo stops /k͡p/ and /kʷ/

and /$\widehat{gb}$/ and /g^w/did not differ significantly from one another in closure duration, but were consistently longer than single articulation stops (p. 457). However, the difference was less obvious in Obolo, indicating a language-specific element.

Ladefoged & Maddieson (1996, p. 363-5) compare spectrograms of bilabials with secondary palatalization and of clusters of a bilabial and a palatal approximant; the most consistent difference was in timing, e.g. in a Russian sequence /p^jo/ the high F_2 frequency of the palatalized stop begins to fall immediately at the onset of voicing, but for the contrasting cluster /pjo/ F_2 frequency remains high for the first 100 ms of the vowel before falling. Thus, the duration and/or the timing of articulations seem to be a prominent cue in differentiating doubly articulated stops from clusters, especially in those languages which have clusters which contrast with double articulation stops (e.g. /pj/ contrasting with /p^j/). However, duration can also be affected by vocalic environment: Hardcastle & Roach (1979, p. 534-534) found that the presence of an adjacent back rounded vowel resulted in a faster articulation of the labial component of /kp/ and /tp/ clusters; the study measured degrees of overlap between the two stops, and found that in only 12% of such cases was the first stop released before the onset of the bilabial closure. This suggests that vocalic environment may affect the timing of multiple articulation stops, but how this interacts with timing constraints resulting from the phonological status of the stops is unclear.

3.2.4 Labialized velar stops and approximants

Formant frequencies The defining acoustic characteristic of labialized velars as described by Ohala & Lorentz (1977, p. 581–3) is the extreme lowering of F_2 frequency caused by constrictions at velocity maxima in the vocal tract. The paper describes locations in the vocal tract, specific to each resonant frequency, in which the local velocities of air particles are higher; constrictions in these places lower the resonant frequency from what would be expected in a uniform tube. For the second formant, velocity maxima are located at the lips and, while the lips constrict, at the velum. Therefore when constrictions are made at both these locations the frequency of the second formant is lowered. Due to this characteristic low F_2 frequency of labialized velars, neighbouring front vowels are strongly affected by coarticulation: effects of the labialized velar approximant on formant transitions have been characterized as a steep rising second formant into a following front vowel (Ladefoged, 2001, p. 53; MacEacher et al., 1996, p. 84; Ladefoged et al., 1996).

VOT Voice onset time is typically longer for labialized velars than for single articulation stops; MacEacher et al. (1996, p. 82–83) measure VOT for voiceless unaspirated stops in Wari' and place them in increasing order /p, t, k, k^w/, with /k^w/ three times longer in VOT than /p/.

Burst spectra The acoustic effects of secondary labiovelarization on fronting of stops in front vowel environments have been described in one previous study. Zue (1976) analyzed English stops in clusters with sonorants /r, l, w/: burst peak frequencies were lower for /k/ and /g/ in all sonorant environments than in non-sonorant environments, with means of 1050 Hz and 980 Hz respectively. Zue did not find any significant effect of vocalic context on burst energy in the clusters and concluded that the following sonorant determines the burst energy, with the following vowel having "little or no influence" (1976, p. 141). Suh (2008) found lowering of overall burst frequencies for labialized velars but did not investigate vowel-specific effects.

3.2.5 Acoustic features

Grave and acute Jakobson et al. (1951) and Jakobson & Halle (1956) identify distinctive features based on the acoustic properties of sounds, a theory that was ultimately superceded by feature systems which rely on articulatory gestures. Within the framework they proposed, sounds in which energy is concentrated in the lower frequency range are classified as 'grave'; if there is more energy in the higher frequency range, the sound is 'acute'. Further evidence for acoustic features comes from research in neuroscience; Bouchard & Chang (2014) and Bouchard et al. (2013) found that the location of sensory-motor cortex activity predicted the acoustic parameters of /CV/ sequences.

Within Jakobson's system labial and velar stops, with low F_2 frequencies at the burst and transition, are both grave. Using acoustic features, in contrast to articulatory features, places labials and velars together in a natural class and so provides a potential rationale for perceptual sound change accounts (Ohala, 1989, 1993). Jakobson & Waugh (1979, p. 116–117) point to the typological co-occurrence of labiovelarization as additional evidence in favour of their being placed together in a single natural class.

Flat and sharp In Jakobson's original formulation, lip-rounding (and also retroflexion and pharyngealization) is associated with the feature 'flat', which is itself characterized by lowering of some or all formants (Jakobson et al., 1951, p. 31–32). Flatness may be combined with gravity or acuteness, in which case it contrasts with 'plain'. It is opposed by 'sharp', which manifests as a rise in formant frequency. Secondary labialization, velarization, and pharyngealization are categorized as flat due to their lowering effect on formants; thus, labialized velar sounds are both grave and flat.

Implications The features flat and grave are occasionally cited as explanations for the cross-linguistic tendency for languages to have labialized velars in their stop inventories, but not labialized coronals (e.g. Maddieson & Disner (1984)). There is no obvious articulatory restriction preventing the tongue combining both velar or coronal articulations with lip-rounding, but acoustic features provide a possible explanation for the imbalance in that the combination of a velar constriction with lip-rounding enhances the feature grave (Engstrand et al., 1998; Jakobson et al., 1951). In contrast, dental or alveolar sounds are acute, or [-grave], so the addition of lip-rounding may result in an auditorily ambiguous output. A similar argument is applied to the cross-linguistic absence of both pharyngealization and labialization as secondary articulations; Jakobson et al. 1951, p. 31 analyze them as variant realizations of a single auditorily contrastive feature.

3.2.6 Conclusions

Relative to other obstruents, there is a lack of literature describing the acoustic and articulatory properties of labialized velars in general. Where consonants with velarity and labiality are studied, the labial-velar approximant tends more commonly to be the focus of research. Results of such research are applicable to this study, but there are likely to be salient differences in the articulation and acoustics of stops with secondary labiality that would make it beneficial to provide a more focussed description of labialized velar stops in this chapter.

A more comprehensive study of the acoustics of labialized velars in different vocalic environments than is available in the current literature is therefore needed. It will inform further questions on the sound changes outlined in Chapter 2. Analysis of formant frequency transitions, as well as stop bursts, is necessary to describe not only the general characteristics of labialized velars but how they affect and are affected by neighbouring vowels.

3.2.7 Predictions

It is predicted that the labialized velars analysed in this study will have a longer voice onset time relative to the other stops tested. This is due firstly to their velar place of articulation which, as discussed in §3.2.1, typically causes a longer VOT. Similarly, lip rounding correlates to longer voice onset times. The mass of air inside the vocal tract in front of the closure affects VOT by virtue of its inertia, which impedes the release of air and thus the change of pressure upon release; thus it may be expected that the increased restriction on the release of this mass of air caused by the rounding of the lips and the resulting narrowing of the vocal tract will contribute further to increased VOT for labialized velars. If this is the case then greater rounding of the lips, or greater duration of lip rounding, caused by coarticulation with a following rounded vowel would be expected to increase VOT of labialized velars (and indeed the other stops tested) due to the further increased restriction of the flow of air from the vocal tract.

Ohala (1989) and Garrett & Johnson (2013) discuss the possibility of acoustic similarity between /kw/ and /p/ motivating cross-linguistically common sound changes of /kw/ into /p/. It is expected that labials will have shorter VOT's than labialized velars. Previous research suggests that burst spectra may differ in that labials tend to have diffuse spectra, whilst labialized velars have a prominent low frequency peak. Formant transitions may be acoustically similar due to the effects of lip-rounding in lowering the first and second formant frequencies.

Lowering of F_2 in most vocalic contexts is expected to be observed, as is lowering of F_3 as a result of lip rounding. Although there has been a great deal of research into the contextual effects of different vowels on velar place of articulation, there is little in the literature to suggest whether or not the coarticulatory tongue placement is manifested for velars with contrastive secondary labialization in the same or a similar way as for non-labialized velars. Given the requirement for tongue retraction to make a labial-velar constriction in these sounds it is expected that the same degree of coarticulatory tongue fronting observed for plain velars before front vowels will not be observed for /kw/. The Greek sound changes outlined in Chapter 2 indicate that palatalization of labialized velars is a potential motivation for sound change; acoustic evidence of coarticulatory tongue fronting in labialized velars would provide grounds for an account involving progressive palatalization of labialized velars in front vowel environments in Greek.

The philological evidence presented in Chapter 2 indicated a cross-linguistic tendency for labialized velars to lose contrastive labialization in positions adjacent to rounded vowels. It is predicted that the acoustic study will reveal a high degree of acoustic similarity between contrastively labialized and non-contrastively labialized velar stops in this vocalic environment.

3.3 Acoustic analysis of labialized velars

The literature review in §3.2 reveals a dearth of fine-grained phonetic analysis of labialized velar sounds. The current section will describe the acoustic properties of these sounds in detail in British English and Zapotec, allowing crosslinguistic comparison that will form the basis for hypotheses relating to the sound changes in ancient Greek. As the purpose of the thesis is to investigate the possibility of acoustic motivations for developments of labialized velars to velar, labial, and coronal stops, the acoustic characteristics of all four kinds of stop in specific vocalic environments will be analyzed and compared. Examples of labialized velars from the British National Corpus (Coleman et al., 2012), from material recorded in a studio at the Phonetics Laboratory in Oxford, and from the Zapotec and Chatino

Survey (Sicoli & Kaufman, 2010) will be used to give a detailed description of the acoustics of labialized velars in different vocalic environments. Inclusion of speech gathered “in the wild” will allow for comparison of spontaneous and citation-style speech, in the hope of exposing characteristics of labialized velars that may not be apparent in citation-style recordings alone, and will provide examples of English and Zapotec labialized velars as they are realized acoustically in normal everyday speech. The material recorded in the studio and the Zapotec material will permit comparison of multiple tokens from the same speakers, which is not possible with the BNC data; although the BNC has many speakers, multiple repetitions of the same word from the same speaker are rare.

3.4 Experiment 1: British English Corpus analysis

3.4.1 The corpus

The Audio British National Corpus (Coleman et al., 2012) consists of approximately 6.9 million words of spontaneous English speech. Recordings were collected during the period 1991 to 1994 using analogue audio casette tapes. Participants used personal stereos to record spontaneous natural conversations in non-studio environments. The tapes were digitized in 2009 to 2010 and downsampled to 16 kHz with 16-bit resolution. Orthographic transcriptions, originally produced by hand, were aligned with the audio using a forced aligner based on HTK, with phone transcriptions available as a by-product of forced alignment. The aligned transcriptions are stored in Praat TextGrid format and in simple list format. The corpus is thus searchable by Arpabet phone and by word, with each entry including the tape reference code and start and end times within that tape for each word or phone. Due to the provenance of the recordings the quality of the audio is sometimes not of a sufficient standard for the extraction of very reliable acoustic measurements. As well as this, part of the aim of the original BNC project was to gather examples of real-life speech from a range of regional British English dialects, so there is no dialect consistency in the tokens analysed here; the resulting acoustic data is exemplary of the range of variation across speakers and dialects of British English.

3.4.2 Methodology

Selection of tokens Tokens were extracted for analysis by searching for instances of the required search terms (for the labialized velar tokens, word-initial instances of ‘qu’ in the word dictionary; see Table 3.1 for a list of the ten most frequently attested ‘qu’ words in the corpus. For the comparison phones, minimal pairs were selected where possible to match the most frequently occurring labialized velar initial words. The start and end times for each word, provided by the forced alignment program, were used to play each selection with a buffer of 0.1 s to allow the researcher to make a judgement on 1) whether or not the selection contained appropriate audio, to filter out words beginning with the letter sequence ‘qu’ that did not contain labialized velars, such as *quiche*, and to exclude tokens with poor quality of recording or tokens in which speech was obscured by background noise such as doorbells, barking dogs etc.; 2) quality of alignment, i.e. did the selection actually contain what the transcription indicates; and 3) sex of the speaker. Selections which seemed to be well-aligned were given a tag specifying male or female, copied, and stored as individual .wav files. Poorly aligned or bad quality tokens were discarded. The search program yielded approximately 8000 potential tokens, of which 4808 were suitable for retention and analysis. The 0.1 s buffer added by the selection program to the beginning and end of each selection was then removed using the *trim* function in *sox* (Bagwell, 1998). This left .wav files containing individual words starting from points within or before the closure stage of the initial stop. Files were checked at random and found to have a good standard of alignment accuracy, i.e. the clip contained the target sounds and started during the closure phase of the stop or in a period of silence before the release if the word was utterance initial.

Gender imbalance Although care was taken at the time of recording the corpus to select a balanced sample in terms of speaker age and sex, tokens selected for this research were categorized as male at a rate consistently higher than that of female. This imbalance was consistent for many of the words selected for use in the study. Participants were often recorded speaking in mixed-gender groups, indicating there could be a sociolinguistic explanation for the imbalance. Due to the relatively small number of female tokens, analysis was restricted to male speakers.

Table 3.1: The ten most frequent attestations of English words beginning with /kw/ in the audio British National Corpus (BNC).

Word	Count
quite	1947
question(-s, -ed)	994
quick(-ly, -er)	393
quality	294
quarter(-s, -ly)	243
quote(-s, -ed)	157
quarr(-y, -ies)	136
quid	132
queen	66
quiet	53

3.4.3 Formant frequency analysis

Formant measurements Formant frequencies were extracted due to their importance as acoustic cues to stop identification (Dorman et al., 1977). Formant frequencies were measured using ESPS (Entropic Research Laboratories, 1996). To isolate the vocalic portions of the relevant words, the ESPS function `get_f0` was used to calculate the probability of voicing (expressed as 0 or 1) at intervals of 5 ms throughout the clip. The voicing probability scores produced by `get_f0` were used to isolate the first voiced interval in each .wav file, equivalent to the first vocoid of the word, including the voiced portion of the [w] if it was voiced. An algorithm was developed to resolve issues caused by poor alignment; if the first measurement of voicing probability was 1, indicating that the recording began with voicing—the result of poor alignment causing a previously uttered word to be included in the clip—this interval would be passed over in favour of the second voiced interval, which would be the target vowel. If, on the other hand, the first voicing probability measurement was 0, indicating no voicing was present, it was assumed that the clip was well aligned and lack of voicing indicated the closure phase of the target consonant, so the first voiced interval after that voiceless interval was assumed to be the target vocoid. The methodology was tested on approximately 5% of tokens and this algorithm was found to be successful in negating potential problems caused by poor alignment and in finding and measuring the correct intervals. Formant frequencies from the same 5 ms frames within the voiced intervals were then measured using the `formant` function in ESPS. All the speakers were male (see §3.4.2) so five formants were extracted using an LPC order of 16 and a pre-emphasis constant of 0.9. A high pre-emphasis constant was chosen due to the low signal-to-noise ratio of many tokens, resulting from the recording methods described in §3.4.1. This method was broadly successful but suffered due to the poor audio quality of some tokens; despite the high pre-emphasis constant, interference in the form of broadband noise or buzzing in a number of tokens led to inaccurate formant tracking, which manifests as outliers.

3.4.4 Results for stops preceding front vowels

The tokens English has two high front vowels, /ɪ/ and /i/. Labialized velars preceding the English high front lax vowel /ɪ/ were represented by the words *quick* and *quickly*. For comparative purposes data were also gathered from words which began with three other stops, /p, t, k, and which formed minimal pairs with the /kw/ words. These were *pick* ($n = 74$) and *kick* ($n = 104$). Due to the low number of *tick* samples in the corpus, tokens of *tip* were included with the data for *tick*, giving a total count of 60 tokens. For the /i/ words, the onset stops were represented by a minimal triplet of *queen* ($n = 66$), *penis* ($n = 13$) *teen* ($n = 29$) and *keen* ($n = 73$).

Formant frequency data Mean F_2 and F_3 frequencies and standard deviations for /ɪ/ words are presented in Tables 3.2 and 3.3. Mean F_2 and F_3 frequencies and standard deviations for /i/ words are presented in Tables 3.4 and 3.5.

Boxplots Box and whisker plots for frequencies of the first to third formants for each word are presented in Fig. 3.1 for /ɪ/, and in Fig. 3.2 for /i/ . One box and whisker plot is presented for each timepoint, at intervals of 0.005 s from 0 s after the onset of the vocoid until 0.05 s (measured as described in §3.4.3), showing the dynamic trajectory of each formant.

Fig. 3.1: F_1, F_2, and F_3 measurements for words with vowel /ɪ/ beginning [p, t, k, k^w], taken at intervals of 0.005 s from the onset of voicing.

(a) 'Pick', $n = 74$ **(b)** 'Tip, tick', $n = 60$ **(c)** 'Kick', $n = 104$ **(d)** 'Quick(ly)', $n = 280$

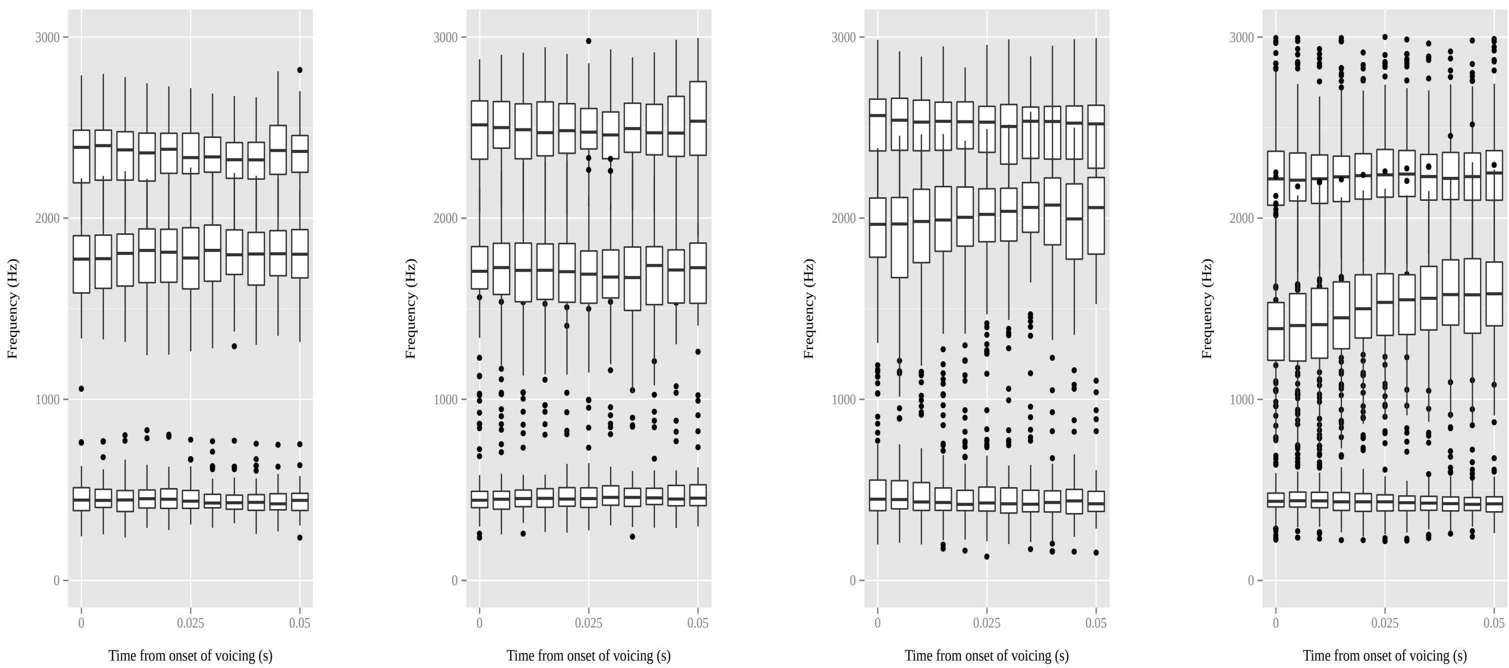


Fig. 3.2: F_1, F_2, and F_3 measurements for words with vowel /i/ beginning [p, t, k, kʷ], taken at intervals of 0.005 s from the onset of voicing.

(a) ‘Penis’, $n = 13$ **(b)** ‘Teen’, $n = 29$ **(c)** ‘Keen’, $n = 73$ **(d)** ‘Queen’, $n = 66$

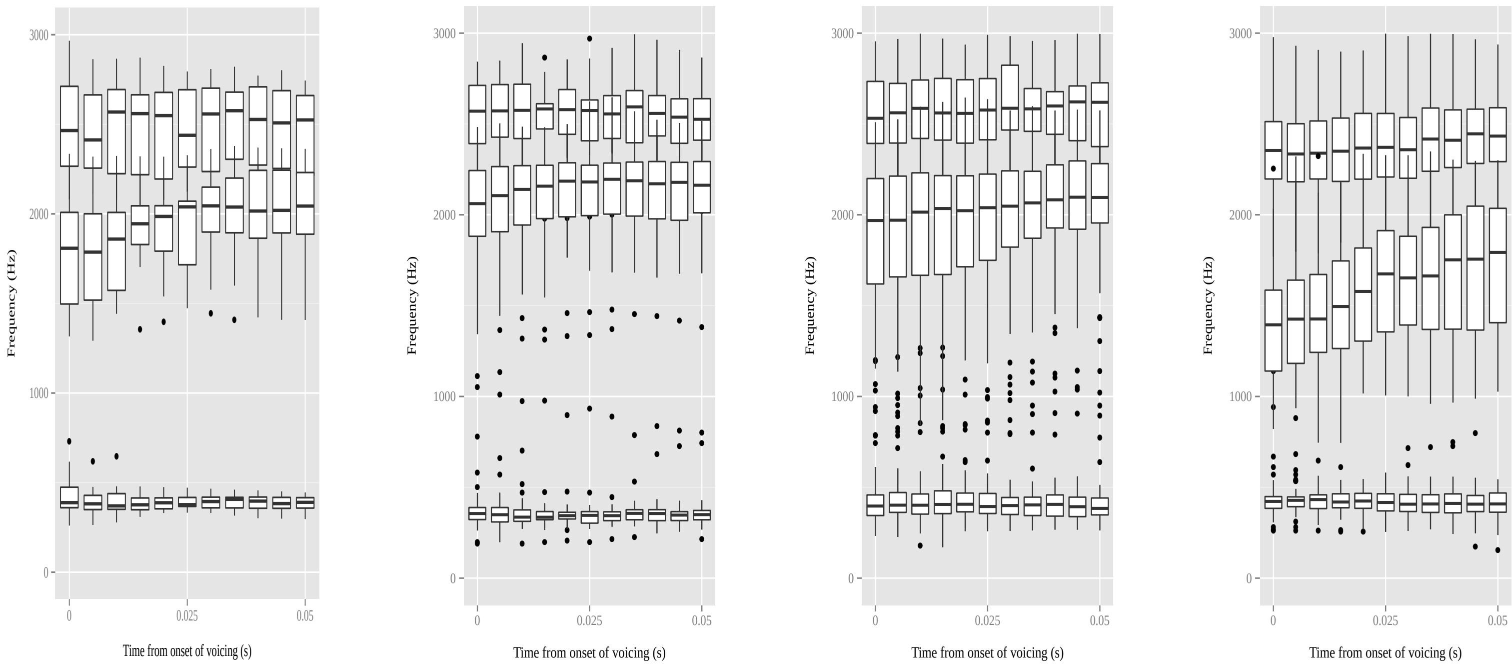


Table 3.2: $\bar{F}_2$ and standard deviations for /ɪ/ tokens (Hz) at 0 ms after the onset of voicing, 0.05 s after the onset of voicing, and for all timepoints from 0-0.05 s of voicing.

Token	$\bar{F}_2$ at 0 ms	*s. d.*	$\bar{F}_2$ at 50 ms	*s. d.*	$\bar{F}_2$ 0-50 ms	*s. d.*
pick	1742	201	1801	211	1782	213
tip, tick	1671	279	1658	255	1676	275
kick	1900	328	1967	360	1937	342
quick(ly)	1380	274	1564	306	1476	294

Table 3.3: $\bar{F}_3$ and standard deviations for /ɪ/ tokens (Hz) at 0 ms after the onset of voicing, 0.05 s after the onset of voicing, and for all timepoints from 0-0.05 s of voicing.

Token	$\bar{F}_3$ at 0 ms	*s. d.*	$\bar{F}_3$ at 50 ms	*s. d.*	$\bar{F}_3$ 0-50 ms	*s. d.*
pick	2359	200	2383	219	2349	191
tip, tick	2490	301	2546	390	2498	295
kick	2522	244	2525	286	2508	230
quick(ly)	2250	304	2267	258	2261	283

Table 3.4: $\bar{F}_2$ and standard deviations for [i] tokens (Hz) at 0 ms after the onset of voicing, 0.05 s after the onset of voicing, and for all timepoints from 0-0.05 s of voicing.

Token	$\bar{F}_2$ at 0 ms	*s. d.*	$\bar{F}_2$ at 50 ms	*s. d.*	$\bar{F}_2$ 0-50 ms	*s. d.*
penis	1774	320	1981	347	1915	83
teen	1946	402	2105	273	2071	322
keen	1898	376	2058	369	1967	384
queen	1392	312	1743	369	1580	356

Table 3.5: $\bar{F}_3$ and standard deviations for [i] tokens (Hz) at 0 ms after the onset of voicing, 0.05 s after the onset of voicing, and for all timepoints from 0-0.05 s of voicing.

Token	$\bar{F}_3$ at 0 ms	*s. d.*	$\bar{F}_3$ at 50 ms	*s. d.*	$\bar{F}_3$ 0-50 ms	*s. d.*
penis	2482	283	2462	244	2476	258
keen	2556	264	2634	277	2605	268
teen	2581	272	2554	226	2588	249
queen	2395	309	2468	281	2416	291

3.4.5 Discussion

Contextually conditioned differences in vowel quality At the onset of voicing, F_1 frequencies are significantly lower for /kwɪ/ words than for /tɪ/ words ($\bar{F}_1$ for /kwɪ/ is 458 Hz, $\bar{F}_1$ for /tɪ/ is 483 Hz; independent two-group t test, $p < 0.001$) but do not differ significantly from /kɪ/ or /pɪ/ frequencies (Fig. 3.1).

F_2 frequencies were significantly lower for /kwɪ/ words than for all other words (see Table 3.2 for $\bar{F}_2$ values; independent two-group t tests, $p < 0.001$). For the /i/ tokens, F_1 is very significantly lower for /ti/ than for /kwi/ and /ki/ (independent two-group t tests, $p < 0.001$) whilst $\bar{F}_1$ for /ki/ and /kwi differs only marginally (independent two-group t tests, $p < 0.05$). F_2 frequency lowering after /kw/ is observable in Table 3.2 in which the distribution of *teen* vowels overlaps entirely in the F_1—F_2 vowel space with that of /ki/ vowels, whilst /kwi/ tokens overlap only partially with /ki/ and the /kwi/ mean lies outside the 95% confidence interval of ti tokens. In the F_2–F_3 vowel space, /kwɪ/ and /pɪ/ cluster closer together in the F_3 dimension whilst /tɪ/ and /kɪ/ cluster higher; there is no significant difference in F_3 means between /kɪ/ and /tɪ/, whilst the other word means do differ significantly). For /i/ words, /ki/ and /ti/ tokens overlap in both dimensions whilst /kwi/ tokens tend towards lower F_2 values.

Contextually conditioned differences in vowel formant transitions The data presented show a strong acoustic dissimilarity between the formant transitions of front vowels after voiceless aspirated labialized velars and front vowels after voiceless aspirated coronals and voiceless aspirated velars. After a coronal stop, F_2 was noticeably and very significantly higher than after a labialized velar (Table 3.4, independent two-group t-test, $p < 0.001$). This is expected at the onset of voicing due to the continued presence of lip rounding after the burst release, but in fact the F_2 frequency of the vowel in *queen* never reaches a height equivalent to that of the vowel in *teen* in the first 50 ms of voicing (Table 3.4, $p < 0.001$). It is not until approximately 80 ms after the onset of voicing that the probability of obtaining a similar result if the *queen* F_2 samples were drawn from the same population as the *teen* F_2 samples exceeds the significance threshold of 5%, indicating a process of gradual convergence ($\bar{F}_2$ of *queen* at 8 ms = 1830 Hz, $\bar{F}_2$ of *teen* = 1993 Hz, independent two-group t-test, $p = 0.06$). Differences in F_3 frequency were similarly significant during the first 50 ms of voicing (Table 3.4, $p < 0.001$) but converged at a similar timepoint. It is possible that, due to the methodology (which used the probability of voicing generated by ESPS `get_f0` to locate each vowel, §3.4.3) formant frequencies from the end of the measured period are more strongly affected by the final voiced nasal than at the start of the vocoid.

After a non-labialized velar stop, F_2 frequencies for /i/ were higher than after both labialized velars and coronal stops throughout the first 50 ms of voicing. The difference between *queen* and *keen* was highly significant at 0 ms, at 50 ms, and over the entire time period (independent two-group t-tests, $p < 0.001$). For *keen* and *teen*, the difference in F_2 across the first 50 ms was small but significant ($\bar{F}_2$ of *keen* = 1967 Hz, $\bar{F}_2$ of *teen* = 2071 Hz giving a difference of 104 Hz, $p < 0.001$). However, for *kick* and *tip/tick*, $\bar{F}_2$ was higher for *kick* than for *tip/tick*.

Dynamic distinctiveness in /kwɪ/ and /kwi/ tokens Mean F_2 frequency after /t/ at the onset of voicing is significantly higher for /i/ ($\bar{F}_2 = 1946$ Hz) than for /ɪ/ ($\bar{F}_2 = 1671$ Hz, $p < 0.01$) whilst after /kw/ it remains similar ($\bar{F}_2 = 1415$ Hz for /kwi/ tokens, $\bar{F}_2$= 1380 Hz for /kwɪ/ tokens, $p > 0.05$); see Table 3.2. This suggests that the acoustic differences between /kw/ and /t/ are even more detectable before /i/ than before /ɪ/, due to the greater difference between the relative lowness of the first formant measurements after /kw/ and the relative height of the /i/ vowel. This evidence suggests that an initially low F_2 which rises sharply during the articulation of the first part of the vowel, coupled with a relatively low F_3, lends a strongly distinctive acoustic characteristic to /kw/ in positions preceding front vowels in British English. This could inhibit perceptual confusion, undermining a hypothesis that acoustic similarity could prompt changes of /kw/ to /t/ in this vocalic context.

/ɪ/ **contexts** When formant transitions into /ɪ/ are compared in all consonantal contexts, it is clear that the characteristically low second formant of /kwɪ/ persists throughout the duration of the vowel; /ɪ/ after /kw/ does not reach a second formant frequency value equivalent to that of /ɪ/ after /t/ at any point before the coda stop. Likewise, the third formant shows consistently lower frequency values for /ɪ/ in positions following /kw/ than after the other stops.

/i/ **contexts** The formant trajectory boxplots for tokens containing /i/ show that the rise in F_2 frequency during /i/ after /kw/ is even greater than for /ɪ/ (Fig. 3.2). For the *queen* tokens, F_2 frequencies showed a mean rise of 351 Hz over the time interval from the onset of voicing, where $\bar{F}_2$= 1392 Hz, to 0.05 s after the onset of voicing, where $\bar{F}_2$= 1743 Hz; the difference in $\bar{F}_2$ at each timepoint was statistically significant, $p < 0.001$. For the *quick(ly)* tokens $\bar{F}_2$ rose by only 173 Hz.

Similarity between /k/ **and** /t/ Figures 3.1 and 3.2 show that there is overlap between the ranges of variation of /k/ and /t/ in both front vowel contexts, despite a very significant difference in the means of F_2 frequency (independent samples t-test, $p < 0.001$). The $\bar{F}_2$ of /i/ after /t/ throughout the voiced period is 2071 Hz, which is within one standard deviation of $\bar{F}_2$ of /i/ after /k/ at 1967 Hz (Table 3.4). The effect for /ɪ/ after /t/ and /k/ is similar, although the two variants of /ɪ/ are slightly less close to one another within the vowel space (Table 3.2). The ability of the human auditory and perceptual systems to distinguish categorically between finely differentiated frequency distributions may mean that this difference is used as a cue to distinguishing /k/ from /t/ in this context, but closely overlapping variation may also present opportunities for effective acoustic convergence; this could in turn lead to acoustically-motivated perceptual confusion. To summarise, /ki/ and /ti/ are possibly confusable, which could lead to historic merger, but /kwi/

and /ti/ and /kwɪ/ and /tɪ/ are not expected to merge on acoustic grounds.

Syllable coda effects In the /ɪ/ set of words, *kick* vowels have a higher F_2 frequency (Fig. 3.1c) than *tip, tick* vowels (Fig. 3.1b), but this pattern is reversed for the /i/ set, in which *teen* vowels (Fig. 3.2b) have a higher F_2 frequency than *keen* vowels (Fig. 3.2c). This can most likely be attributed to anticipatory coarticulatory effects of the syllable coda. For *teen*, the tongue maintains a forward position near the alveolar ridge for both stops, resulting in a more forward position for the vowel and therefore higher F_2 frequency; this is not the case for *tip, tick* words in which the tongue either retracts to produce a velar constriction or retracts to rest during the labial closure. *Kick* vowels have a particularly high F_2 frequency, possibly due to increased palatalization by coarticulation to both an initial and final palatal. In all cases, mean F_2 frequencies are higher for /k, t/ words than for /kw/ words, regardless of syllable coda.

3.4.6 Results for stops preceding back round vowels

The tokens The most frequently occurring words in which labialized velars occurred preceding back rounded vowels were *quality* (n = 108) and *quarter* (n = 180). The word *politics* was selected to provide a comparison of labialized velars and bilabial plosives in a phonetic context as similar as possible to *quality*. There were no good words or word combinations in sufficient numbers to enable an effective comparison of *quality* with initial /t/ or /k/. Of the 243 instances of *quarter* identified in the BNC, 180 tokens were extracted and analysed. Comparable words were sought beginning with /p, t, k/ to allow comparison (see Table 3.6). To retain maximum comparability in the phonetic enviornment, tokens were selected in which the target onset stop was followed by a vowel equivalent to the open back vowel in *quarter*, followed by a coronal stop /t/ or glottal stop (both variants of /t/ were accepted as otherwise the number of tokens would have been too small to provide useful data) followed, if possible, by a mid central vowel /ə/. Due to the limited number of appropriate /p/ tokens, words beginning with *port* were also accepted—although the majority matched the pattern for *quarter*, i.e. 'porter', 'portable'—and *caught the* and *caught her* were accepted as well as *caught a* and *court of*. *Caught* and *court* were homophonous for nearly all tokens, based on native speaker judgements.

Table 3.6: Token counts for the back round vowel /ɔ/ environment.

/kw/	/p/	/t/	/k/
quarter (180)	porter (2)	taught (53)	caught a (6)
	portable (19)		caught her (2)
	port (12)		caught the (6)
			court of (20)
180	33	53	34

Acoustics of the vowels

Formant frequency data The mean formant frequencies for *quality* and *politics* set are presented in Table 3.7. Mean formant frequencies from the *quarter*, *caught/court*, and *port(-)* tokens are presented in Table 3.8.

Boxplots Box and whisker plots for frequencies of the first to third formants are presented in Fig. 3.4 for *quarter*, *caught*, *port* tokens and in Fig. 3.3 for *quality*, *politics* tokens. One box and whisker plot is presented for each timepoint from 0 ms after the onset of the vowel until 0.05 s (measured as described in §3.4.3), showing the dynamic trajectory of each formant.

Table 3.7: $\bar{F}_2$, $\bar{F}_3$, and standard deviations for *politics* and *quality* tokens (Hz) at 0 ms after the onset of voicing, 50 ms after the onset of voicing, and aggregated for time points across the first 50 ms of voicing.

	$\bar{F}_2$ at 0 ms	*s. d.*	$\bar{F}_2$ at 5 ms	s. d.	$\bar{F}_2$ 0-5 ms	*s. d.*
politics	1012	321	948	199	965	242
quality	1013	328	960	189	982	277
	$\bar{F}_3$ at 0 ms	*s. d.*	$\bar{F}_3$ at 5 ms	s. d.	$\bar{F}_3$ 0-5 ms	*s. d.*
politics	2531	256	2512	278	2524	274
quality	2387	330	2416	327	2405	326

Table 3.8: $\bar{F}_2$, $\bar{F}_3$ and standard deviations for *quarter*, *port(-)*, *taught*, and *caught/court*, tokens (Hz) at 0 ms after the onset of voicing, 5 ms after the onset of voicing, and aggregated for time points across the first 5 ms of voicing.

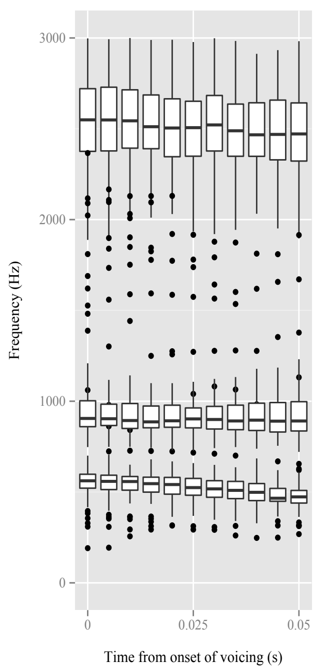
	$\bar{F}_2$ at 0 ms	*s. d.*	$\bar{F}_2$ at 5 ms	s. d.	$\bar{F}_2$ 0-5 ms	*s. d.*
port(-)	1248	619	1137	652	1190	628
taught	1198	394	1076	367	1138	386
caught	1055	362	970	383	1005	370
quarter	1109	489	1045	433	1073	469
	$\bar{F}_3$ at 0 ms	*s. d.*	$\bar{F}_3$ at 5 ms	s. d.	$\bar{F}_3$ 0-5 ms	*s. d.*
port(-)	2616	381	2679	387	2658	371
taught	2441	280	2463	262	2462	254
caught	2344	357	2434	278	2415	297
quarter	2476	358	2473	369	2477	347

Table 3.9: Means and p values (derived using Welch two sample t tests) for BNC material, for acoustic comparison of tokens in the /kwo/ and /po/ categories. Data from male speakers.

At the onset of voicing			
	Quality	*Politics*	p value
F_1 (Hz)	529	568	0.08
F_2 (Hz)	1013	1012	0.97
F_3 (Hz)	2387	2531	0.001**
During first 0.05 s after the onset of voicing			
	Quality	*Politics*	p value
F_1 (Hz)	528	529	0.72
F_2 (Hz)	982	965	0.16
F_3 (Hz)	2405	2524	<0.0001***
At the onset of voicing			
	Quarter	*Porter*	p value
F_1 (Hz)	445	480	0.34
F_2 (Hz)	1109	1248	0.23
F_3 (Hz)	2476	2616	0.058
During first 0.05 s after the onset of voicing			
	Quarter	*Porter*	p value
F_1 (Hz)	442	460	0.03*
F_2 (Hz)	1073	1190	0.003**
F_3 (Hz)	2477	2658	<0.0001***

Fig. 3.3: F_1, F_2, and F_3 values for *quality* ($n = 108$) and *politics* ($n = 75$)

(a) 'Politics', $n = 75$



(b) 'Quality', $n = 108$

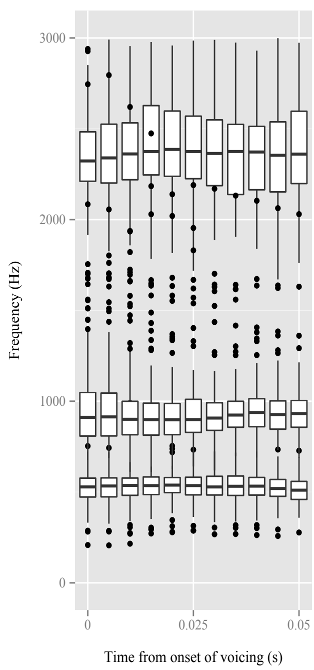
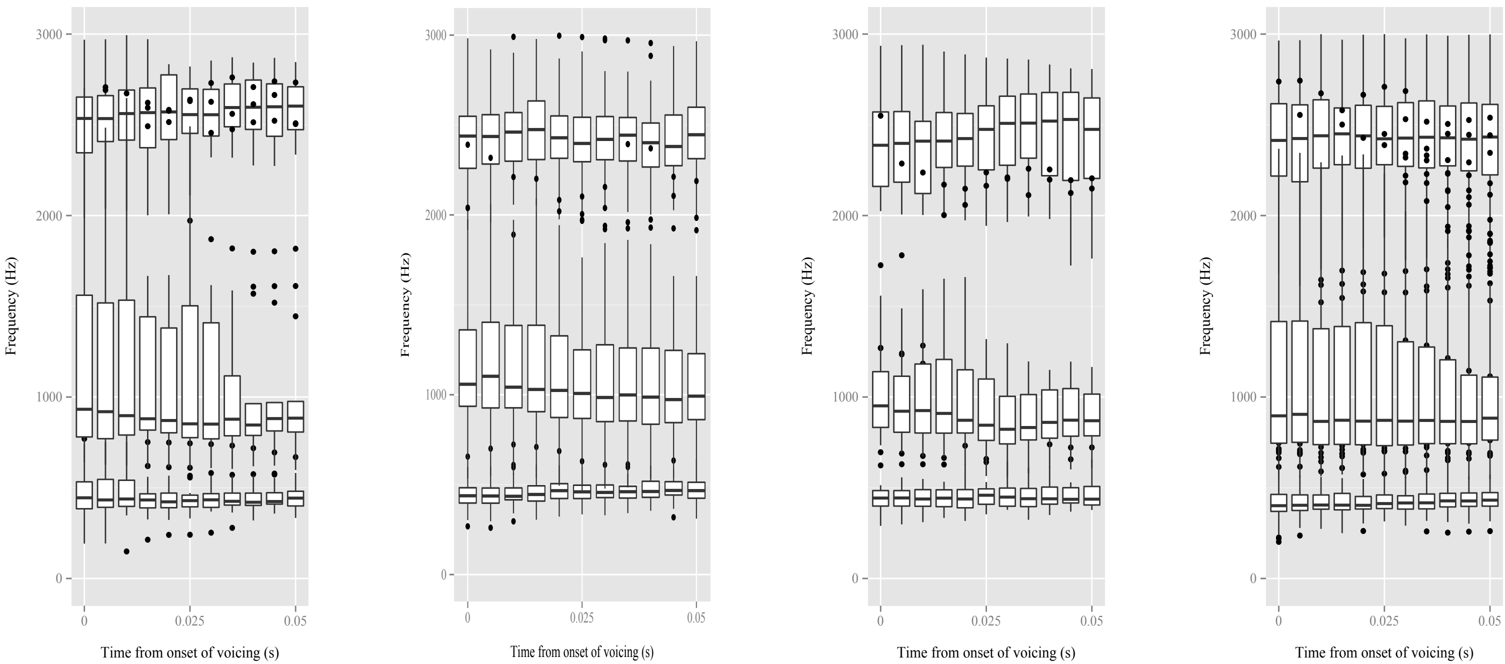


Fig. 3.4: F_1, F_2, and F_3 values for *quarter*, *port*, *taught*, and *caught/court* words.

(a) 'Port', $n = 33$ **(b)** 'Taught', $n = 53$ **(c)** 'Caught, court', $n = 34$ **(d)** 'Quarter', $n = 180$



3.4.7 Discussion

Contextually conditioned differences in vowel formant transitions Lip rounding is expected to be present throughout the initial stop and vowel of *quality* and *politics* due to anticipatory coarticulation to the roundedness of the vowel. The acoustic effects of lip rounding as a contrastive feature of vowel articulation are, primarily, lowering of the first three formants (Ladefoged, 2001, p. 160 ff.). This was confirmed in the results: F_1, F_2, and F_3 frequencies are low in both contexts. The difference in F_2 between *quality* and *politics* is significant for the voiced interval overall (for *quality*, $\bar{F}_2 = 1043$ Hz over the first 0.05 s of voicing, for *politics* $\bar{F}_2 = 944$ Hz, $p < 0.001$) although the mean frequencies are statistically indistinguishable at the first time point (values given in Table 3.9). The difference in F_3, however, is significant at every time point sampled over the duration of the voiced interval (for *quality* $\bar{F}_3 = 2342$ Hz, for *politics* $\bar{F}_3 = 2516$ Hz, $p < 0.001$). This suggests that the relative lowering of F_3, resulting from increased lip-rounding of the vowel by coarticulation with the labialized velar, is potentially a more salient cue to labialized velarity than that of F_2.

Data for the *quarter*, *caught/court*, *port(-)* set show expected similarities to the *quality–politics* set (see Figure 3.4). Formant frequency trajectories of F_1 and F_2 are similar for the *quarter* and *caught/court* tokens. The biggest difference between tokens in the *caught/court* group and tokens in the *quarter* group is to be found in the frequency of F_2, which remains significantly lower for *caught/court* than for *quarter* throughout voicing ($\bar{F}_2$ for *caught/court* $= 1005$ Hz, $\bar{F}_2$ for *quarter* $= 1073$, independent two group t test, $p < 0.005$). However, the difference is only 68 Hz, so whilst it is significant it may not be auditorily perceptible. F_2 frequencies of *caught/court* and *port(-)* tokens are similar. As in the *quality/politics* set, the most striking acoustic difference in the formant frequencies of the back round vowel is the lowering of F_3 under the influence of lip rounding of the labialized velar, observed for *quarter* but not for the *port(-)* set ($\bar{F}_3$ for *quarter* $= 2477$ Hz, $\bar{F}_3$ for *port(-)* $=$

2658 Hz, independent two groups t test, $p < 0.001$).

Role of F_3 The evidence indicates that before back rounded vowels the potentially biggest cues to the distinction between /kw/ and /p/ may be the lowering of the F_3 frequencies from the onset of voicing; for the /kwɔ/ and /kwɒ/ category words, F_3 was significantly lower than for the /pɔ/ and /pɒ/ category words, whilst F_1 and F_2 did not differ between the two groups (Table 3.9). Differences in vowel quality between *quality–politics* and *quarter–porter* do not appear to manifest themselves in significant F_3 differences. For the vowels of *quality* and *politics*, F_3 frequencies were similar to those of—respectively—*quarter* and *port(-)* at vocalic onset. However, *quality* and *politics* had higher F_1 and lower F_2 frequencies than *quarter* and *port(-)*. This suggests that the lowering of F_3 in the transition into the vowel is indicative of the quality of the preceding stop or cluster rather than of the quality of the vowel.

3.4.8 Burst spectra analysis

Spectral measurements Log power was calculated at intervals of 160 samples (with a frame length of 200 samples) throughout each clip using ESPS functions `frame` and `pwr` (Entropic Research Laboratories, 1996). Voicing probabilities were calculated using ESPS `get_f0`. The burst release was identified by finding the time at which 1) the probability of voicing was 0, and 2) the power plot reached a peak (peaks were defined as points at which the sum of a power measurement and the following power measurement switched from positive to negative). This methodology was automated and tested, and found to be successful in locating the burst in 24 out of 25 randomly selected test tokens (constituting 3.5% of the total number of tokens); the one failure was due to poor alignment which resulted in the burst being cut off the start of the recording. This rate of accuracy was considered acceptable as the high n made available using the automated methodology should compensate for a certain amount of “noise” in the data. The burst times were taken as time points for extraction of spectral slices using the ESPS functions `refcof` and `me_spec`, with 512 FFT points, a Hamming window, and an LPC order of 20.

Emphasis Intensity of human speech decreases at approximately 6 dB per octave (Stevens, 1998, p. 69). The spectra were emphasized by +6 dB per octave to counteract this effect after extraction and increase the visibility of peaks in the higher frequency range.

Calculating means Mean intensity and 95% confidence intervals at each frequency were calculated using the R function `summarySE` as defined by Chang (2012, p. 363).

Root mean square Root mean squares (RMS) were calculated to give the average intensity across the entire frequency range for each spectrum. RMS values are derived using equation (3.1), in which x is a range of intensity measurements in dB and $n = 256$.

$$x_{rms} = \sqrt{\frac{1}{n}(x_1^2 + x_2^2 + \ldots x_n^2)} \qquad (3.1)$$

The RMS calculation was implemented using the R programming language.

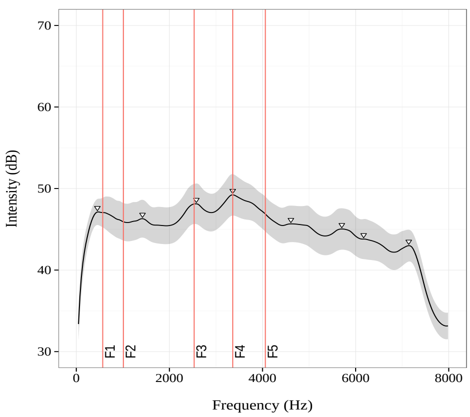
Spectral peaks The difference in intensity (on the y-axis) between each frequency measurement point (x-axis) and the following frequency measurement point was calculated throughout each spectrum. Spectral peaks were then identified as points at which the differentials switched from negative to positive, i.e. as a rising slope reverses its trajectory and starts to fall.

Formant measurements Formant frequencies were measured as described in §3.4.3. Some tokens were lost due to difficulties in formant tracking on lower quality recordings, and due to inconsistencies in the voicing probability patterns caused by slight misalignment in the BNC; thus n is lower for the formant frequencies than for the spectral peak measurements.

3.4.9 Results

Fig. 3.5: Mean LPC spectra at the burst for British English tokens of *quality* and *politics*, multiple speakers. Shaded areas indicate 95% confidence interval. Arrows indicate spectral peaks. Vertical lines show mean formant frequencies measured at the point of voicing onset for each CV vocoid.

(a) /pɒ/ burst of *politics*, with $\bar{F}_1$–$\bar{F}_5$ of /ɒ/, n = 75; RMS = 45.7 dB



(b) /kwɒ/ burst of *quality*, with $\bar{F}_1$–$\bar{F}_5$ of /ɒ/ , n = 115; RMS = 49.1 dB

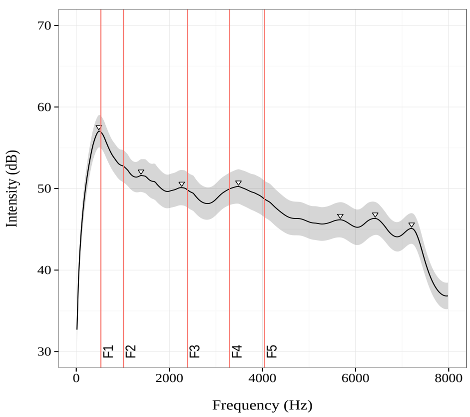
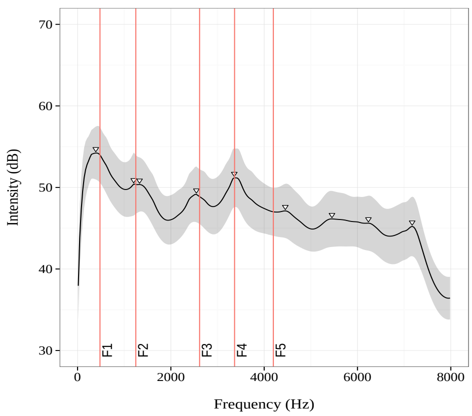
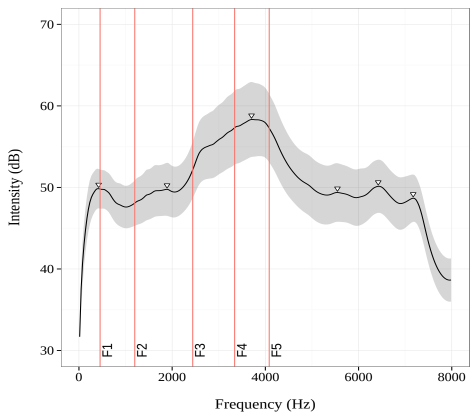


Fig. 3.6: Mean LPC spectra at the burst for British English CV clusters with vowel /ɔ/ for contrastively labialized velar stops, non-contrastively labialized velar stops, and labial stops; multiple speakers. Shaded areas indicate 95% confidence interval. Arrows indicate spectral peaks. Vertical lines show mean formant frequencies measured at the point of voicing onset for each CV vocoid.

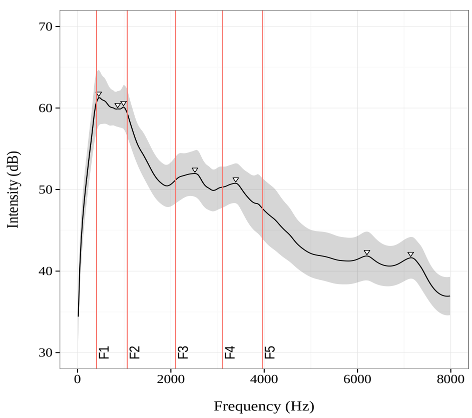
(a) /pɔ/ burst of *port(er)*, with $\bar{F}_1$–$\bar{F}_5$ of /ɔ/, n = 33; RMS = 47.7 dB



(b) /tɔ/ burst of *taught*, with $\bar{F}_1$–$\bar{F}_5$ of /ɔ/, n = 52; RMS = 51.9 dB



(c) /kɔ/ burst of *caught*/*court*, with $\bar{F}_1$–$\bar{F}_5$ /ɔ/, n = 32; RMS = 48.3 dB



(d) /kwɔ/ burst of *quarter*, with $\bar{F}_1$–$\bar{F}_5$ of /ɔ/, n = 171; RMS = 49.7 dB

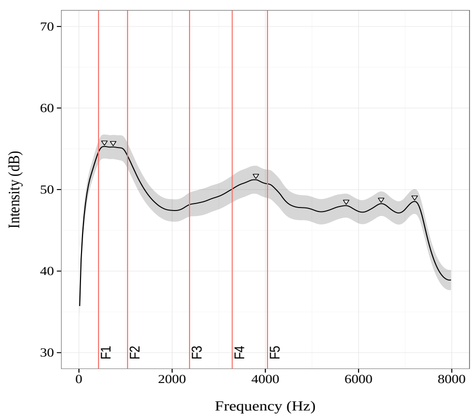
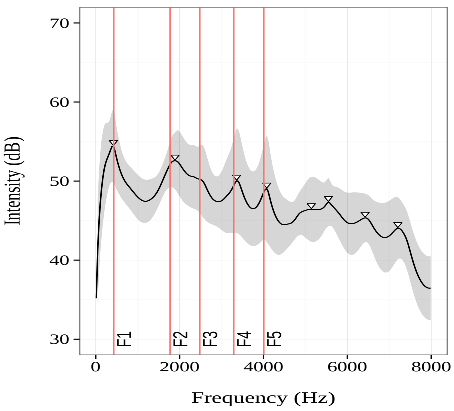
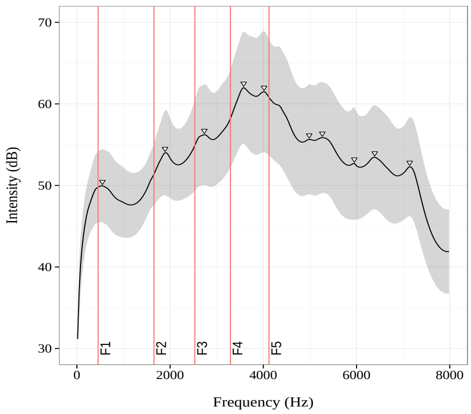


Fig. 3.7: Mean LPC spectra at the burst for British English CV clusters with vowel /i/ for contrastively labialized velar stops, labial stops, non-labialized velar stops, and coronal stops; multiple speakers. Shaded areas indicate 95% confidence interval. Arrows indicate spectral peaks. Vertical lines show mean formant frequencies measured at the point of voicing onset for each CV vocoid.

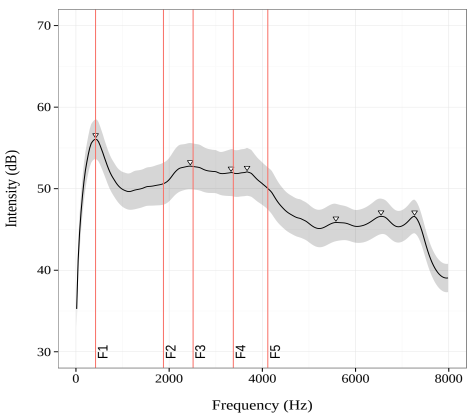
(a) /pi/ burst of *penis*, with $\bar{F}_1$–$\bar{F}_5$ of /i/, $n = 13$; RMS = 47.3 dB



(b) /ti/ burst of *teen*, with $\bar{F}_1$–$\bar{F}_5$ of /i/, $n = 31$; RMS = 54.7 dB



(c) /ki/ burst of *keen*, with $\bar{F}_1$–$\bar{F}_5$ of /i/, $n = 73$; RMS = 49.7 dB



(d) /kwi/ burst of *queen*, with $\bar{F}_1$–$\bar{F}_5$ of /i/, $n = 66$; RMS = 51.5 dB

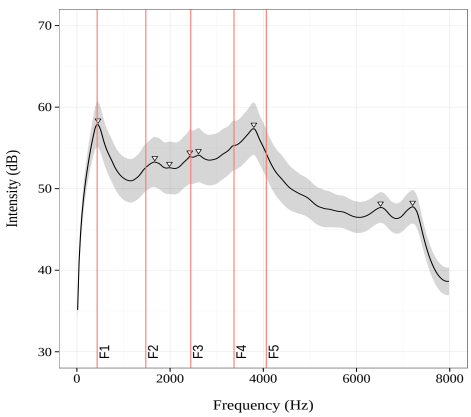
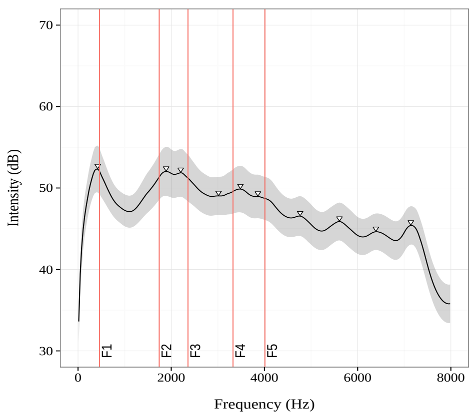
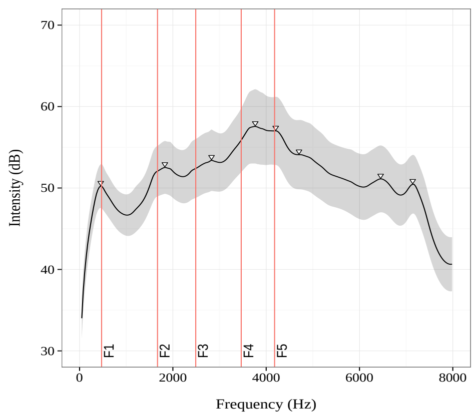


Fig. 3.8: Mean LPC spectra at the burst for British English CV clusters with vowel /ɪ/ for contrastively labialized velar stops, non-contrastively labialized velar stops, coronal stops, and labial stops; multiple speakers. Shaded areas indicate 95% confidence interval. Arrows indicate spectral peaks. Vertical lines show mean formant frequencies measured at the point of voicing onset for each CV vocoid.

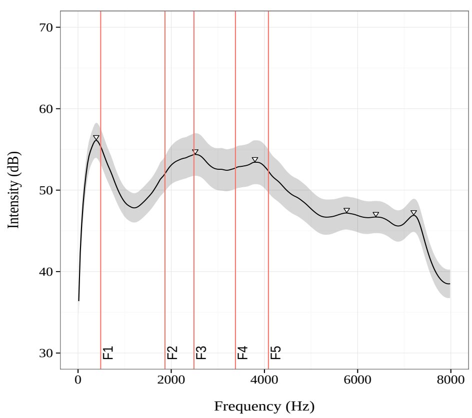
(a) /pɪ/ burst of *pick*, with $\bar{F}_1$–$\bar{F}_5$ of /ɪ/ $n = 74$; RMS = 48 dB



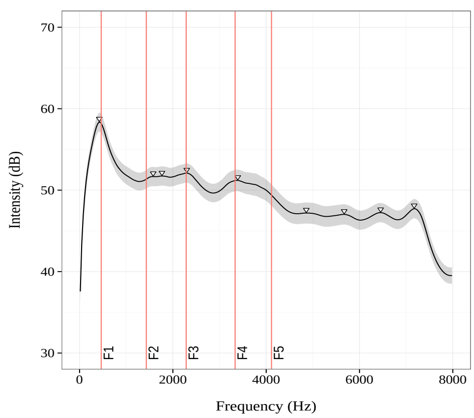
(b) /tɪ/ burst of *tip/tick*, with $\bar{F}_1$–$\bar{F}_5$ of /ɪ/ $n = 60$; RMS = 52.9 dB



(c) /kɪ/ burst of *kick*, with $\bar{F}_1$–$\bar{F}_5$ of /ɪ/, $n = 101$; RMS = 50.6 dB



(d) /kwɪ/ burst of *quick(-)*, with $\bar{F}_1$–$\bar{F}_5$ of /ɪ/, $n = 277$; RMS = 50.1 dB



3.4.10 Discussion

Labial bursts The spectra for /pɒ/ and /pɔ/ are diffuse, with peaks at F_1, F_3, and F_4 of the following vowel for both spectra; the /pɔ/ spectrum also features a peak at F_2 of the following vowel (Figs. 3.5a and 3.6a). Before front vowels /ɪ/ and /i/, the labial burst is diffuse, with multiple peaks at or around the frequencies of each formant (Figures 3.7a and 3.8a).

Coronal bursts Before /ɒ/, /t/ bursts have a prominent peak between F_4 and F_5 of the following vowel (Fig. 3.6b). Before both front vowels, energy is heavily concentrated in the high frequency range for coronal bursts with multiple peaks and a prominent peak above F_4 of the following vowel. The prominent peak is more intense for /i/ (Fig. 3.7b) than for /ɪ/ (Fig. 3.8b).

Non-labialized velar bursts The mean spectrum for the /kɔ/ bursts is compact, with energy concentrated in the lower frequency range and a dominant peak at 484 Hz; there are additional lower energy peaks at 2547 Hz and 3422 Hz (Fig. 3.6c). F_2 of the following vocoid is relatively low, with a frequency of 1065 Hz. The spectra for /kɪ/ and /ki/ are relatively diffuse, and F_2 frequency of the following vocoid is higher for /kɪ/ (1900 Hz) and for /ki/ (1898 Hz) than for /kɔ/ (1055 Hz). The /kɪ/ burst has a spectral peak at 391 Hz, equivalent to the frequency of F_1 of the following vowel, and additional peaks at 2516 Hz and 3799 Hz; the higher frequency peaks are of roughly equivalent intensity to the F_1 peak with a range of 53.4 to 56.0 dB.

The /ki/ burst spectrum has a prominent peak at 422 Hz, equivalent to the F_1 frequency of the following vocoid, and high frequency peaks at 2453 Hz, 3328 Hz and 3672 Hz (Fig. 3.7c). These peaks are all located between F_2 and F_5 of the following front vowel (F_2 frequency of /i/ = 1865 Hz, F_5 frequency = 4086 Hz; Table 3.4). Intensity for the three higher frequency peaks is consistent across the peaks (ranging from 51.9 dB to 52.7 dB) whilst the low frequency peak has a higher intensity of 56 dB.

Labialized velar bursts The spectrum for /kwɔ/ is quite diffuse, with prominent peaks at 578 Hz and 785 Hz (Fig. 3.6d). These values are close to F_1 of the following vowel, 446 Hz (Table 3.8). There is a secondary peak at 3828 Hz.

The burst spectrum for /kwɪ/ has a prominent peak at 453 Hz and no equally prominent high frequency peaks; there are some low intensity peaks in the range 1609 Hz to 3422 Hz, but overall the spectrum is compact (Fig. 3.8d).

The burst spectrum for /kwi/ is very diffuse, with two prominent peaks at 484 Hz and 3828 Hz and several other high intensity peaks in the range between F_2 and F_4 of the following vowel (Fig. 3.7d).

Root mean squares In all vocalic contexts the average intensity across the spectral burst, calculated using the RMS equation as outlined in equation 3.1, increased in the order /p, k, kw, t/, i.e. /p/ had the lowest average intensity at the burst and /t/ the highest.

3.4.11 Conclusions

The data suggests that an initially low F_2 which rises sharply during the articulation of the vowel, coupled with a relatively low F_3, lends a strongly distinctive acoustic characteristic to the formant transitions of front vowels following /kw/. The data suggests that a greater degree of frontness in the context vowel is required for /kw/ bursts to show acoustic evidence of velar fronting.

The /kwɪ/ and /kwɔ/ spectra are more similar to one another than to /kwi/, which more closely resembles the spectra of non-labialized velars before front vowels i.e. /kɪ/ and /ki/. Although formant transitions for /kɔ/ and /kwɔ/ for these tokens are statistically similar (Figs. 3.4c and 3.4d) there is a clear difference in energy distribution at the burst (Figs. 3.6c and 3.6d): the burst spectrum in /kwɔ/ has a series of prominent peaks above F_5 of the following vowel which are not observed in the /kɔ/ spectrum.

All spectra show a peak just below the frequency of the vowel's F_5. This peak varies in intensity and is most prominent in the /kwi/, /ki/ and /kɪ/ spectra and least prominent in the /kwɪ/, /kwɔ/ and /kɔ/ spectra. This could be caused by frication energy resulting from multiple release of the velar or increased area of tongue contact for velar stops.

In the data presented here, instances of /i/ are significantly more front than instances of /ɪ/ (using F_2 height as a measure of frontness). This corresponds to increased burst energy represented by prominent peaks in the frequency range between F_2 and F_4 of a following vowel for /k/ tokens, confirming Keating & Lahiri (1993)'s results, but also increased burst energy within that frequency range for /kw/ tokens. Bursts before /i/ have more intense peaks in this region that bursts before /ɪ/, indicating a proportional relationship between vowel frontness and labialized velar fronting.

However, F_2 frequency for both front vowels after /kw/ is lower than after /k/. Mean F_2 frequency for /ɪ/ after /k/ is 1893 Hz, and 1511 Hz after /kw/; mean F_2 frequency of /i/ after /k/ is 1945 Hz, and 1717 Hz after /kw/ (Figs. 3.7d and 3.8d). These differences are both highly significant (independent samples t test, $p < 0.001$) and imply that formant transitions could remain a highly useful acoustic cue to differentiation between labialized and non-labialized stops in front vowel environments, despite observed similarities in burst energy distribution. Interestingly, this is a contrast to the examples with /ɔ/, in which the burst is markedly dissimilar for /kw/ than for /k/, but formant transitions are not significantly different.

3.5 Experiment 2: British English studio recording analysis

3.5.1 Introduction

Speech was recorded at the Phonetics Laboratory in Oxford for analysis in this acoustic study and for use as stimuli in Experiment 4 (see Chapter 4). The materials from the BNC were not of a sufficient quality to use as tokens in a perception experiment so it was necessary to record additional material that does not suffer from the recording quality problems found in the BNC (e.g. background noise, distance from microphone etc.). The acoustics of that material are presented in this section, allowing a nuanced investigation of the relationship between acoustics and perception in the perception experiment that will be described in Chapter 4.

3.5.2 Methodology

Selection of tokens Speech produced by six native Southern British English speakers was recorded. Due to the aims and requirements of Experiment 4, stimulus words containing the onset segments /p, t, k, kw/ and nuclei /ɛ, ei/, /ɪ/, and /ɔ, ɒ/ were recorded. There are twelve possible combinations of these obstruent and vowel sounds (Table 3.10).

Table 3.10: Possible obstruent and vowel combinations.

	p	t	k	kw
e	pe	te	ke	kwe
o	po	to	ko	kwo
i	pi	ti	ki	kwi

Sets of monosyllabic, minimally contrasting English words were identified and divided into subsets instantiating sequences of relevant stops and vowels. For ease of reference in this paper these groups of words will be referred to by the '*CV*' label under which they are subdivided in Table 3.10 (e.g. /te/ refers to *test*, *take*). Due to a paucity of appropriate words in English, it was necessary for the /pe/, /te/, /ke/, /kwe/ groups to be represented by words with [eɪ] or [ɛ], and the /po/, /to/, /ko/, /kwo/ groups by words containing [ɒ] or [ɔ].

Table 3.11: Minimal triplets and, where possible, quadruplets representing each combination of obstruent and vowel.

	p	t	k	kʷ
e	*pest*	*test*	–	*quest*
	[pʰɛst]	[tʰɛst]	–	[kʷʰɛst]
	–	*take*	*cake*	*quake*
	–	[tʰeik]	[kʰeik]	[kʷʰeik]
o	*porn*	*torn*	*corn*	*quorn*
	[pʰɔn]	[tʰɔn]	[kʰɔn]	[kʷʰwɔn]
	pod	*tod*	*cod*	*quad*
	[pʰɒd]	[tʰɒd]	[kʰɒd]	[kʷʰwɒd]
i	*pill*	*till*	*kill*	*quill*
	[pʰɪl]	[tʰɪl]	[kʰɪl]	[kʰʷwɪl]
	pit	*tit*	*kit*	*quit*
	[pʰɪt]	[tʰɪt]	[kʰɪt]	[kʰʷwɪt]
	pick	*tick*	*kick*	*quick*
	[pʰɪk]	[tʰɪk]	[kʰɪk]	[kʰʷwɪk]
	pip	*tip*	*kip*	*quip*
	[pʰɪp]	[tʰɪp]	[kʰɪp]	[kʰʷwɪp]

Subjects and recording The subjects were three female (referred to as 1F, 2F, and 3F) and three male (1M, 2M, and 3M) speakers of Standard Southern

British English. No speakers reported suffering from hearing loss or speech or language impairments. Each subject was recorded in a sound-insulated studio reading individual words from a computer screen; the subjects read every stimulus and distracter in randomized order, at five second intervals and without a carrier sentence, i.e. they were read as isolated citation forms. The recordings were made directly onto CD (16 bit, 44.1 kHz, on one channel of a stereo recording) using an Audio Technica AT4031 microphone, a Symetrix 302 preamplifier, and an HHB CDR-850 professional compact disc recorder. The recordings were then converted to .wav format for processing; the empty second channel was discarded and the audio was downsampled to 16 kHz.

Segmentation Each stimulus word was segmented into onset stop, vowel, and coda, by hand using TextGrids in Praat, based on visual assessment of the spectrogram and waveform of each stimulus. The onset stop was labelled as the duration from the first burst until the onset of voicing, as defined by the detection of glottal pulses by the software. The vocalic interval was defined as the interval of voicing, not including formants visible due to pre-voicing during aspiration. For words beginning /kw/, the measurement of the duration of the voiced interval included the entire voiced segment, i.e. the segment marked /w/ on Figs. 3.9d and 3.10d plus the following vowel. In English, glides are partially devoiced in positions immediately following a voiceless aspirate, making segmentation into aspirated voiceless stop and glide difficult; since lip-rounding is expected to be present during the velar closure due to anticipatory coarticulation, it is not possible to identify a separate aspiration interval and devoiced glide. The result is that in this study formant measurements for the glide begin at the voiced portion of the glide, whilst the devoiced portion of the glide is captured by spectral slices taken at the burst.

In the case of stimulus words with final /l/ (for most tokens, the speakers produced clear rather than dark /l/ with F_2 around or above 1000 Hz, see Fig. 3.9) the end of the vowel and the onset of the coda was defined as the point at which F_2 and F_3 begin to diverge, as F_2 falls under the influence of tongue backing and F_3 rises as the tongue moves to contact the alveolar ridge. This methodology is somewhat subjective due to the lack of obvious segmental boundaries between adjacent vowels and glides, but was applied consistently across the tokens to ensure as much accuracy as possible. The study is concerned with vowel quality rather than duration, and formant frequencies were not measured at the end of the period defined as the vowel (where coarticulatory effects of following resonants would be most problematic), so the definitions used are sufficient for the purposes of the research. Duration of the stop closure could not be measured as the stimuli were produced in citation style without a carrier sentence.

Formant measurement Mean frequencies of the first three formants were measured at the onset of voicing and the temporal midpoint of the voiced interval in each word (the midpoint was determined by hand using Praat; although vowel formants are present before the onset of voicing (and continue into the lateral for those words which end in laterals) for ease of reference the frequencies measured at the midpoint of voicing will be referred to as 'mid-vocalic').

Voice onset time (VOT) measurement Voice onset times were measured by calculating the duration of the initial obstruent segments in the TextGrids. In most cases, aspiration began immediately following the burst so it was not possible to accurately measure burst duration separately; voice onset time was measured as the duration from the start of the first burst to the onset of voicing as determined during the segmentation process described in §3.5.2, following Lisker & Abramson (1964, p. 389).

3.5.3 Results

Formant frequencies for males and females were analysed separately due to significant differences between the sexes: because of their larger vocal tracts males typically have lower formant frequencies than females and that tendency was manifested in this data. Formant frequencies and standard deviations for both sexes are presented in Table 3.12. Example spectrograms of the /i/ and /ɒ/ conditions are provided in Figures 3.9 and 3.10. Voice onset times are presented in Table 3.13.

Table 3.12: Mean frequencies (Hz) and standard deviations (Hz, in parentheses) of the first three formants of words in each *CV* group, measured at the onset of voicing and at the middle of the vocalic segment.

Female Speakers						
CV Group	$\bar{F}_1$ at voice onset	$\bar{F}_2$ at voice onset	$\bar{F}_3$ at voice onset	$\bar{F}_1$ mid-vocalic	$\bar{F}_2$ mid-vocalic	$\bar{F}_3$ mid-vocalic
pe	929 (119)	2109 (68)	2973 (202)	795 (88)	2040 (180)	2914 (404)
pi	564 (564)	2006 (209)	2485 (226)	555 (56)	2067 (178)	2487 (235)
po	461 (126)	887 (132)	2691 (525)	495 (48)	817 (125)	2677 (500)
te	779 (205)	2132 (127)	2989 (239)	746 (104)	2122 (151)	3112 (306)
ti	481 (183)	1990 (187)	2550 (254)	531 (69)	2077 (140)	2568 (223)
to	442 (152)	982 (173)	2143 (591)	483 (78)	901 (150)	2282(566)
ke	759 (104)	2231 (143)	2930 (329)	688 (83)	2191 (232)	2920 (346)
ki	444 (132)	1944 (185)	2570 (180)	522 (45)	2068 (95)	2553 (200)
ko	470 (121)	912 (160)	2507 (477)	517 (85)	887 (167)	2489 (532)
kwe	516 (174)	1070 (331)	2613 (148)	685 (71)	1543 (249)	2592 (114)
kwi	458 (81)	1343 (240)	2611 (251)	553 (49)	1867 (109)	2647 (201)
kwo	408 (89)	742 (99)	2569 (380)	488 (40)	841 (144)	2631 (530)

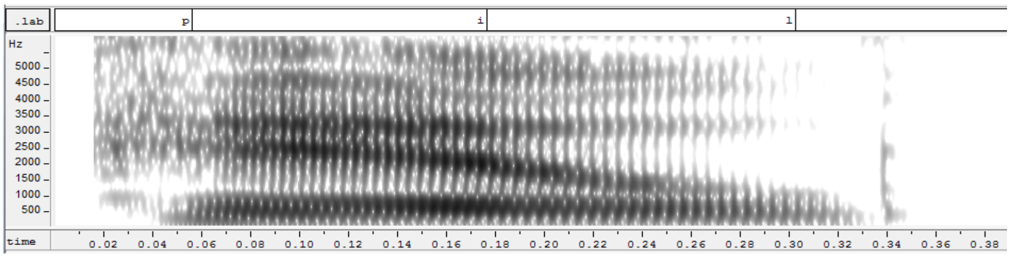
Male Speakers						
CV Group	$\bar{F}_1$ at voice onset	$\bar{F}_2$ at voice onset	$\bar{F}_3$ at voice onset	$\bar{F}_1$ mid-vocalic	$\bar{F}_2$ mid-vocalic	$\bar{F}_3$ mid-vocalic
pe	600 (84)	1725 (71)	2513 (123)	631 (63)	1742 (91)	2583 (222)
pi	416 (91)	1845 (110)	2482 (230)	476 (49)	1810 (161)	2480 (217)
po	434 (160)	739 (149)	2454 (274)	417 (71)	716 (86)	2584 (251)
te	492 (99)	1872 (412)	2583 (139)	562 (79)	1782 (138)	2525 (160)
ti	429 (44)	1848 (129)	2602 (133)	455 (45)	1869 (144)	2561 (131)
to	376 (108)	785 (164)	2415 (257	435 (75)	744)115	2480 (325)
ke	547 (97)	1857 (130)	2643 (305)	534 (73)	1889 (184)	2516 (233)
ki	406(39)	1912 (79)	2506 (257)	447 (33)	1874 (111)	2458 (191)
ko	401 (144)	769 (161)	2325 (224)	430 (78)	762 (81)	2412 (246)
kwe	369 (94)	832 (213)	2221 (110)	530 (72)	1206 (307)	2257 (93)
kwi	356 (77)	998 (360)	2102 (99)	465 (46)	1390 (266)	2147 (56)
kwo	320 (70)	629 (57)	2258 (188)	406 (53)	700 (58)	2384 (234)

Table 3.13: Means (ms) and standard deviations (in parentheses) of voice onset times, by speaker and onset stop.

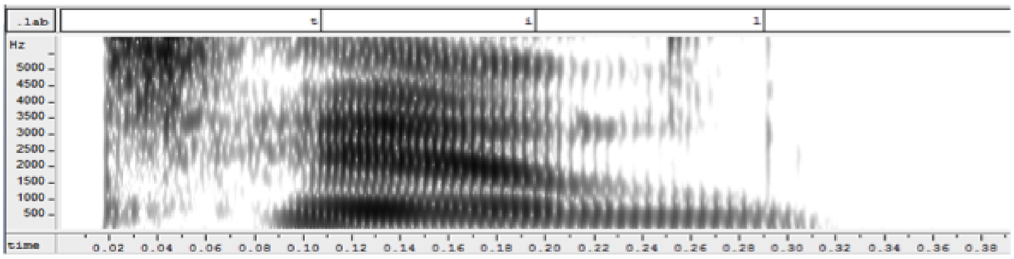
Speaker	p	t	k	kw	Mean (all)
1F	79.6 (12.9)	90.2 (21.6)	95.9 (21.6)	96.3 (18.7)	91.2 (17.0)
2F	72.4 (17.3)	79.5 (13.5)	73.2 (13.9)	101.2 (10.0)	81.7 (17.4)
3F	79.4 (14.7)	90.2 (11.7)	87.5 (14.4)	98.6 (7.9)	89.6 (13.4)
1M	50.0 (16.1)	45.0 (15.6)	66.5 (20.8)	89.5 (21.6)	62.9 (25.4)
2M	91.9 (24.0)	106.6 (25.7)	117.0 (16.8)	129.3 (24.9)	112.0 (26.0)
3M	78.6 (18.4)	80.7 (17.4)	95.4 (14.2)	126.8 (15.5)	96.5 (25.3)
Mean (all)	75.1 (21.1)	82.0 (24.7)	89.3 (23.3)	107.3 (22.9)	89.0 (25.9)

Fig. 3.9: Spectrograms of a female Standard Southern British English speaker saying *pill*, *till*, *kill* and *quill*. The general acoustic similarity of the vowel transitions of *pill*, *till* and *kill* can be seen from their F_2 frequencies, which are high at the onset of voicing and gradually decrease during the transition to the lateral approximant. In contrast, F_2 in *quill* is low at the vocalic onset and rises sharply before decreasing again. One representative sample has been selected for each word. The sampling rate was 16 kHz but the frequency scale on these spectrograms is limited to 6000 Hz in order to more clearly display the lower three formants.

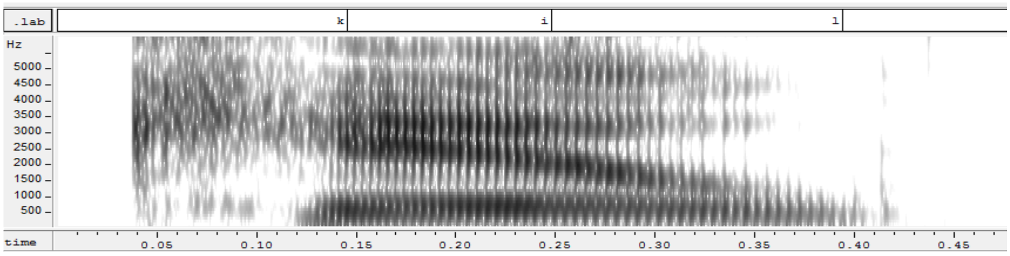
(a) [pʰɪl]



(b) [tʰɪl]



(c) [kʰɪl]



(d) [kʰʷɪl]

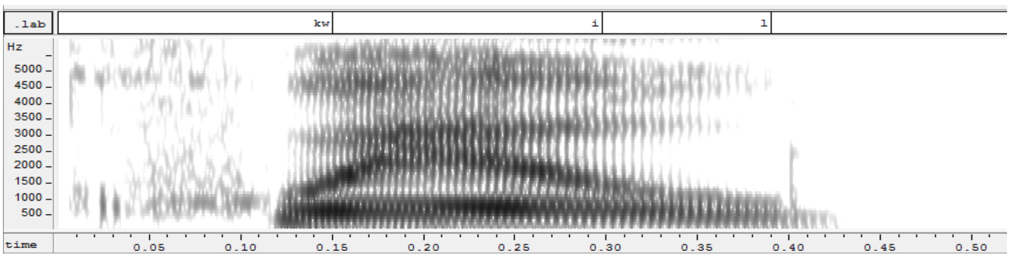
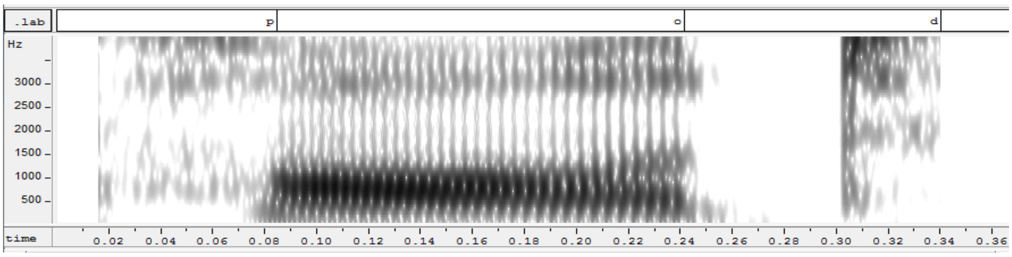
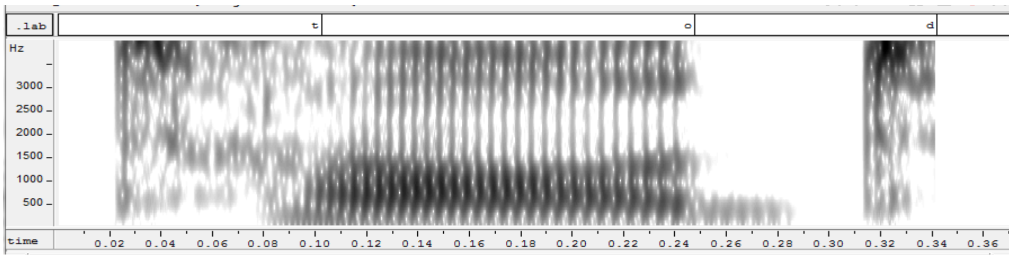


Fig. 3.10: Spectrograms of a female standard Southern British English speaker saying *pod*, *tod*, *cod*, and *quad*. One representative sample has been selected for each word. The sampling rate was 16 kHz but the frequency scale on these spectrograms is limited to 4000 Hz in order to more clearly display the lower three formants (this is smaller than the scale in Figure 3.9 due to the lower F_2 and F_3 frequencies in these ‘ words).

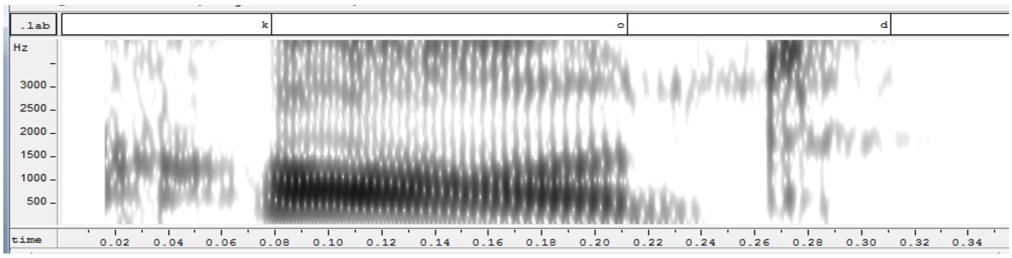
(a) [pʰɒd]



(b) [tʰɒd]



(c) [kʰɒd]



(d) [kʰʷwɒd]

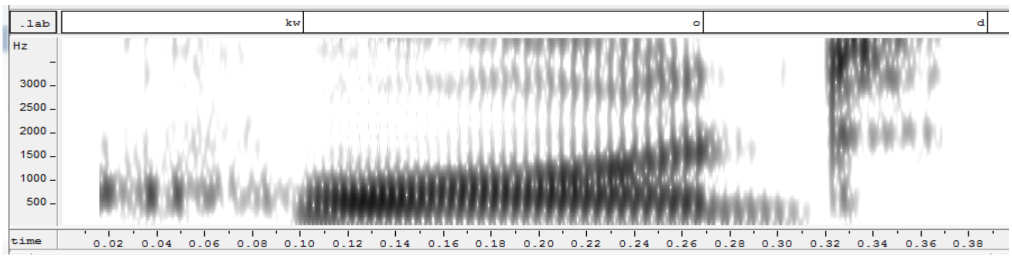


Fig. 3.11: $\bar{F}_1$ vs $\bar{F}_2$ at the onset of voicing (left panels) and mid-vocalically (right panels) in each word group for female(upper panels) and male (lower panels) speakers ($n = 9$ for each data point).

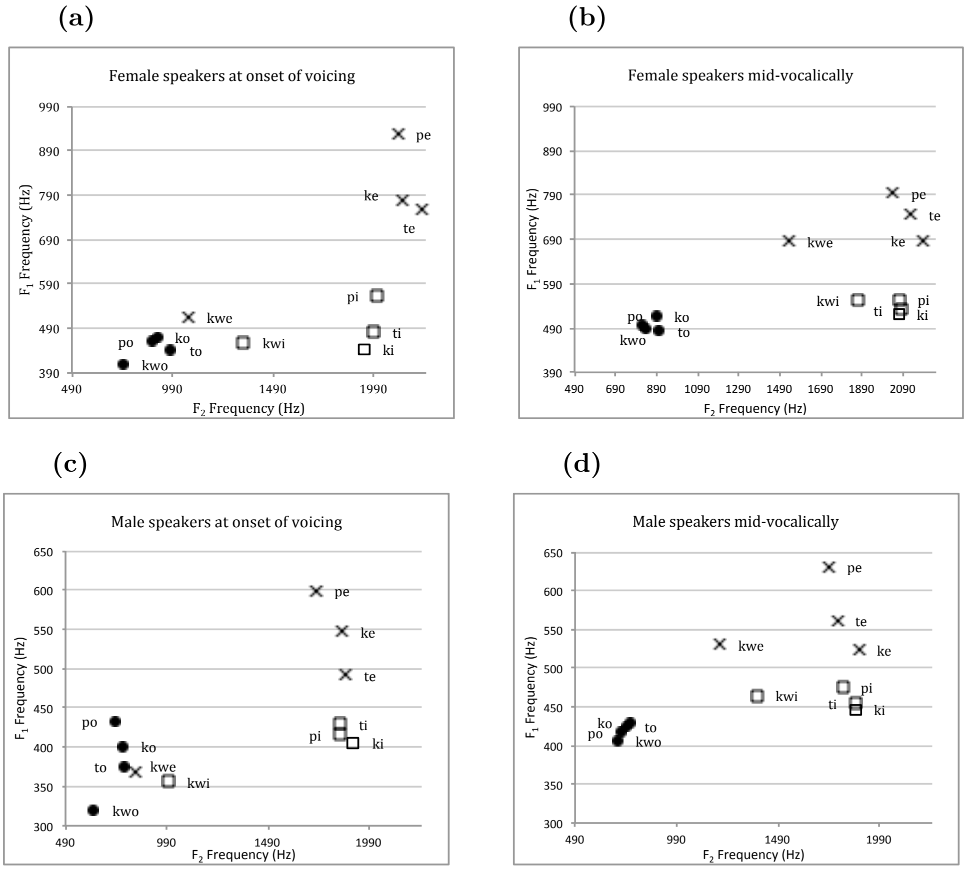


Fig. 3.12: Mean voice onset times by onset stop and vowel.

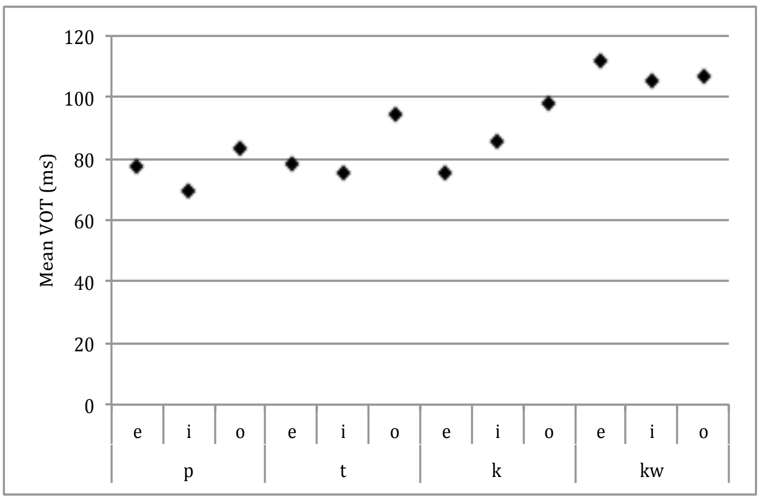


Table 3.14: Means and p values (derived using Welch two sample t tests) for laboratory recorded stimulus material, by sex of speaker. There is one p value less than 0.05, indicating a statistically significant difference between the mean formant frequency values (marked in italics).

Female speakers at onset of voicing			
	/kwo/	/po/	p
F_1 (Hz)	408	461	0.33
F_2 (Hz)	742	887	*0.02**
F_3 (Hz)	2569	2691	0.59
Male speakers at onset of voicing			
	/kwo/	/po/	p
F_1 (Hz)	320	434	0.08
F_2 (Hz)	629	739	0.06
F_3 (Hz)	2258	2454	0.10

3.5.4 Discussion

General observations Words in the /kwo/ group were characterized by low F_1 and F_2 frequencies, and words in the /pi/, /ti/ and /ki/, /kwi/ groups were characterized by relatively high F_2 frequencies both at voice onset and mid-vocalically. F_3 was generally lower for /kwo/ tokens at the onset of voicing.

Vowel quality No significant differences in voice onset time or in formant frequencies at the onset of voicing were found between words with /ɔ/ and words with /ɒ/ for any individual speaker (independent sample t-tests, $p > 0.05$) justifying the anaylsis of words in these categories together under the categories /po/, /to/, /ko/, /kwo/. Likewise, no significant differences were found between the voice onset times or formant frequencies at voice onset of words with /ɛ/ and words with /ɛɪ/ for each onset obstruent (independent sample t-tests, $p > 0.05$). There were no significant differences in formant values for the different words containing /ɪ/ for any onset group (one-way ANOVA, $p < 0.05$). [1]

Vowel duration There were differences in the durations of the vocalic interval for words with /ɛɪ/ and words with /ɛ/. The duration of the vocalic interval did not differ significantly between different onset stops (i.e. the duration of the vowel could not be a cue to the identification of the onset). There were no significant differences in voice onset times for the different words containing /ɪ/ for any onset group (one-way ANOVA, $p > 0.05$).

Formant frequencies The first and second formant frequencies plotted in Figure 3.11 show the main acoustic differences between the twelve CV categories. It has been observed that labialized velars are characterized by relatively low second formant values (Ohala & Lorentz, 1977) and this was borne out by the data collected here. In each vocalic environment labialized velars had lower second formants at the onset of voicing than the other stops tested. The primary difference between /kwo/ and /po/ groups was found to be F_2 frequency, but the difference was only significant for female speakers (Table 3.14).

Bursts and VOT Multiple bursts were observed in the majority of tokens in the /ko/ groups for all speakers (but not in the /ke/ or /ki/ groups) and occurred in most words in the /kwo/, /kwe/ and /kwi/ groups (see Figs. 3.9c, 3.9d, 3.10c and 3.10d for typical examples of /ki/, /kwi/, /ko/ and /kwo/). VOT did not appear to be a salient cue to vowel quality for stops preceding /ɛ/ and /ɪ/ as it did not differ consistently or significantly before these vowels. VOTs were consistently longer

[1]This is not merely an effect of small sample size; for the BNC data studied in Experiment 1, the difference was only verging on significance between [ɔ] and [ɒ] at the onset of voicing in either F_2 or F_3 (for /kw/ words, *quarter* $\bar{F}_2 = 1109$ Hz, *quality* $\bar{F}_2 = 1013$ Hz, $p > 0.05$; *quarter* $\bar{F}_3 =$ 2476 Hz, *quality* $\bar{F}_3 = 2387$ Hz, $p = 0.04$; for /p/ words, *port(-)*, $\bar{F}_2 = 1248$ Hz, *politics* $\bar{F}_2 =$ 1012 Hz, $p = 0.04$; *port* $\bar{F}_3) = 2616$ Hz, *politics* $\bar{F}_3 = 2531$ Hz, $p = 0.25$).

preceding /ɒ, ɔ/ for all word groups except /kw/ (/kwe/ has a longer VOT), but the greater length of VOT before rounded vowels was only significant for /to/ words (independent samples t test comparing mean VOT of /to/ and /ti/ words, $p < 0.05$). Words in the /kwo/ group had the longest VOT in all vocalic environments (Figure 3.12). Ladefoged & Maddieson (1996, p. 333–334) describe clusters of velars and labials as having durations typically 1.5 times to double those of double articulation stops cross-linguistically. It was not possible to measure closure duration in this study due to the citation-style recordings, but the data confirm that labialized velar sounds in this study do have longer aspiration durations than other single articulation stops. However, they do not have aspiration durations that are double the length of the single articulation stops (the mean labialized velar duration of 107.3 ms is 1.3 times that of the other stops, which have a combined mean duration of 82.1 ms).

Speaker-specific differences The laboratory data permits comparison of /kwo/ and /po/ tokens from the same speakers, which was not possible using the BNC data collected in Experiment 1. As in the data collected from the BNC and described in §3.4, frequencies of the first two formants of words in the /kwo/ and /po/ groups of the studio-recorded data were statistically indistinguishable at the onset of voicing; however, unlike in the BNC data, this lack of statistical difference extends into the F_3 frequency values. The explanation for this discrepancy could be found in the difference in sample size; when divided by speaker sex, there were only 9 tokens from 3 speakers in the laboratory recorded data, compared to over one hundred tokens from a much greater number of speakers in the BNC data. Mean formant frequencies at voicing onset and p values for statistical significance of difference are provided in Table 3.15. At the mid-point of the vocalic interval, no significant differences were found between /kwo/ words and /po/ words for any speakers.

Table 3.15: Mean formant frequencies at voicing onset and p values (derived using Welch two sample t tests) for laboratory recorded material, by individual speaker. p values less than 0.05, indicating a statistically significant difference between the mean formant frequencies, are in italics.

Speaker 1F			
	/kwo/	/po/	p value
F_1 (Hz)	492	623	*0.02**
F_2 (Hz	848	1057	*0.006***
F_3 (Hz)	2474	2529	0.84
Speaker 2F			
	/kwo/	/po/	p value
F_1 (Hz)	310	377	0.09
F_2 (Hz)	646	827	*0.016**
F_3 (Hz)	2917	3056	*0.0098***
Speaker 3F			
	/kwo/	/po/	p value
F_1 (Hz)	390	385	0.91
F_2 (Hz)	700	779	0.12
F_3 (Hz)	485	481	0.96

Speaker 1M			
	/kwo/	/po/	p value
F_1 (Hz)	381	517	0.10
F_2 (Hz)	648	833	0.83
F_3 (Hz)	2486	2771	*0.008***
Speaker 2M			
	/kwo/	/po/	p value
F_1 (Hz)	233	352	0.42
F_2 (Hz)	618	649	0.81
F_3 (Hz)	2070	2208	0.37
Speaker 3M			
	/kwo/	/po/	p value
F_1 (Hz)	345	431	0.49
F_2 (Hz)	621	736	0.30
F_3 (Hz)	2217	2383	*0.02**

Fig. 3.13: Mean F_1 vs F_2 values at the onset of voicing for each word grouping, broken down by speaker.

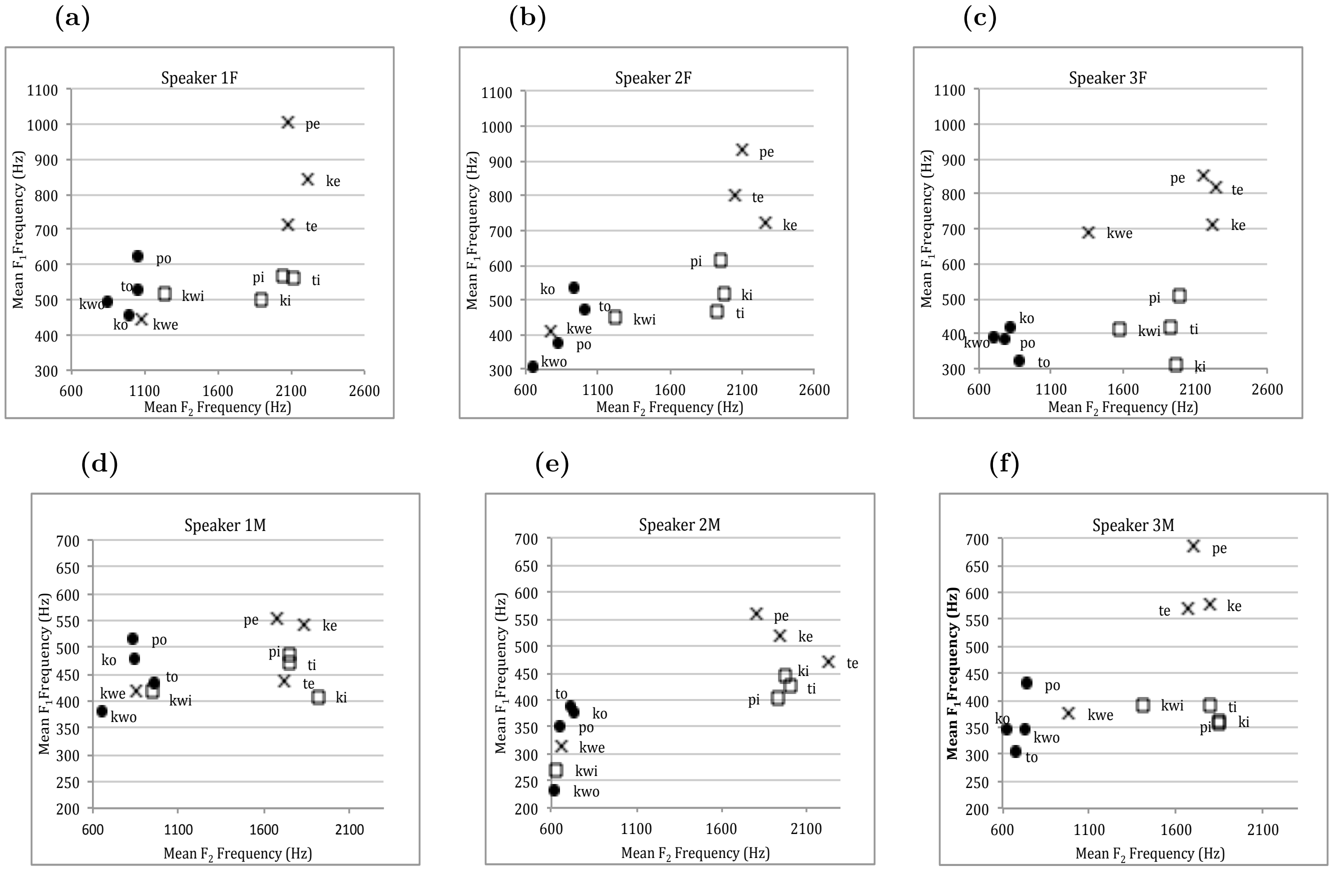
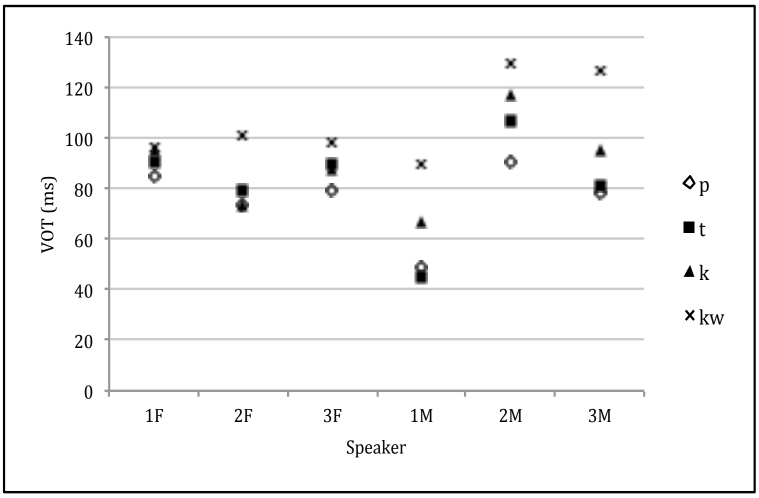


Fig. 3.14: Mean voice onset times for individual speakers, by place of articulation.



VOT Male and female speakers did not differ significantly in VOT. Although actual values differed by speaker, the relative lengths of voice onset time remained broadly similar (Figure 3.14). Tokens in the /kwe/, /kwi/, /kwo/ groups had the longest VOT for all speakers, and those in the /pe/, /pi/, /po/ groups had the shortest for all but one speaker. The fact that the differences in VOT between speakers were in many cases greater than the differences between stops indicates that VOT by itself may not be a particularly salient cue to place.

Formant transitions Whilst sporadic, occasionally significant individual differences in formant frequencies between speakers were observed in all word groups, relative acoustic distances between labialized velars and the outcome obstruents identified in the introduction were largely consistent across speakers at the onset of voicing (Figure 3.13). For speakers 1F, 1M and 3M, words in the /kwo/ and /po/ groups were not more acoustically similar to one another than were words in the /kwo/ and /to/ groups. Words in the /kwo/ and /ko/ groups were acoustically similar for all speakers with the exception of speakers 2F and 2M, for whom /kwo/ and /ko/ were slightly more dissimilar; it is expected that /kwo/ and /ko/ would be spectrally similar at voicing onset, as both have a velar with coarticulatory rounding, [k^{hw}]. A difference between /kwo/ and /ko/ may be more likely to manifest more obviously in the dynamics of the beginning of the vocoid. For the majority of speakers words in the /kwi/ group clustered near the /po/, /to/, /ko/ groups, but the high F_2 frequencies of speaker 3F and 3M's /kwi/ words rendered them slightly more similar to /pi/, /ti/, /ki/ words, although they nevertheless did not cluster near those words. For all speakers /kwe/ words were acoustically dissimilar to /ke/ words in both F_1 and F_2 frequency (although speaker 3F had an exceptionally high mean second formant frequency for /kwe/ words). This is as expected given that /kwe/ has initial [k^{hw}] whereas /ke/ has initial [k̟ʰʲ].

3.5.5 Conclusions

Results for formant transitions confirmed those observed for the BNC corpus data, with labialized velars lowering F_2 of a following vowel. In the F_1–F_2 vowel space, /kwo/ and /po/ tokens were not more similar to one another than to other stops for either male or female speakers, suggesting that they are not acoustically similar in terms of formant transitions. In positions preceding front vowels, Euclidean distances were very high for labialized velars versus the other stops tested when plotted in the F_1–F_2 vowel space, suggesting a high degree of acoustic distinctiveness. This was the case for both front vowels, with /kwe/ in fact more distant from other /Ce/ sequences than /kwɪ/ is from other /Ci/ clusters at onset on the F_1 axis, indicating a stronger formant frequency lowering effect for this vowel.

Voice onset times for labialized velar stops were longer than for the other stops tested for all speakers. Those stops also showed lengthened VOT in back round vowel environments, suggesting that rounding may have a lengthening effect on VOT generally.

3.6 Experiment 3: Western Zapotec Corpus analysis

Western Zapotec Zapotec is spoken by approximately 425,000 speakers, predominantly in Oaxaca, Mexico. In many dialects of Zapotec the inherited labialized velar developed to a bilabial plosive; this is not the case for Western Zapotec, which is the variety discussed in the following sections. Western Zapotec retains labialized velars, although not in positions adjacent to back rounded vowels.

The corpus The Zapotec material used in this study was recorded in 2004-2005 and consists of phrases and sentences elicited by researchers in one-to-one interviews (Sicoli & Kaufman, 2010). Many speakers responded to some questions in Spanish and some questions in Zapotec, meaning the number of tokens elicited from each speaker tends to vary somewhat. The interview method means that a large quantity of speech data was elicited from each speaker, allowing for close comparison of variation dependent upon vowel context within the utterances of each individual speaker.

The corpus is stored as wav files accompanied by transcriptions in ELAN Annotation Format (EAF) transcribed by trained native speakers using a consistent orthography.

Extraction of speech data Three speakers of Western Zapotec were included in this experiment; two are female and one male. Statistically significant differences in female and male fundamental frequencies provide an additional rationale for reporting data from each speaker separately. The speakers are referred to here using the codenames ZPP-PED, ZPP-MAT, and ZPP-ANT as they are described in the corpus (Sicoli & Kaufman, 2010). ZPP-MAT and ZPP-ANT are female, ZPP-PED is male. EAF transcription files for the three speakers of Western Zapotec were converted to Praat TextGrid files, which were then searched for instances of /kw/ and /k/ before the high front vowel /i/, mid front vowel /e/, and back vowel /a/. This gave start times for each sentence response to the interviewer's stimulus question, and the tokens were then extracted from the sentence by hand using Praat, and saved as individual clips. Clips are in the format /VCV/ to retain contextual information, but the preceding vowel has been disregarded for the purposes of this study.

Table 3.16: Number of tokens of each CV for each speaker.

	ZPP-MAT	ZPP-PED	ZPP-ANT
CV	n	n	n
/ka/	37	28	15
/ke/	18	12	5
/ki/	31	12	17
/kwa/	30	5	15
/kwe/	33	25	19
/kwi/	31	12	6

Spectral measurements Plots of power over time were calculated for each clip using Wavesurfer (Sjölander & Beskow, 2000) and the burst release was identified as the timepoint between intervals of voicing at which the power peaked. LPC spectral slices were then taken by hand from this time point using Wavesurfer, with 512 FFT points, a Hamming window, and an LPC order of 20.

Emphasis The tokens were emphasised in the same way described for the British English tokens in §3.4.2.

Calculating means Means were calculated in the same way described for the British English tokens in §3.4.2.

Spectral peaks Spectral peaks were calculated in the same way described for the British English tokens in §3.4.2.

Formant measurements Formant frequencies were measured at intervals of 0.005 s from the beginning to the end of each clip using ESPS function `formant`, following which voicing probabilities generated by `get_f0` were used to find the voiced portions by matching to the pattern /VCV/ within the clip.

3.6.1 Results

Spectral data for the burst releases are presented in Figures 3.15, 3.17, and 3.19. Formant frequency measurements are summarized in Table 3.17; boxplots are presented in Figures 3.16, 3.18, and 3.20.

Fig. 3.15: Mean LPC spectra at the burst for Zapotec CV clusters of /kw, k/ with vowels /a, e, i/, speaker ZPP-MAT. Shaded areas indicate 95% confidence interval. Arrows indicate spectral peaks. Vertical lines show mean formant frequencies measured at the point of voicing onset for each CV vocoid.

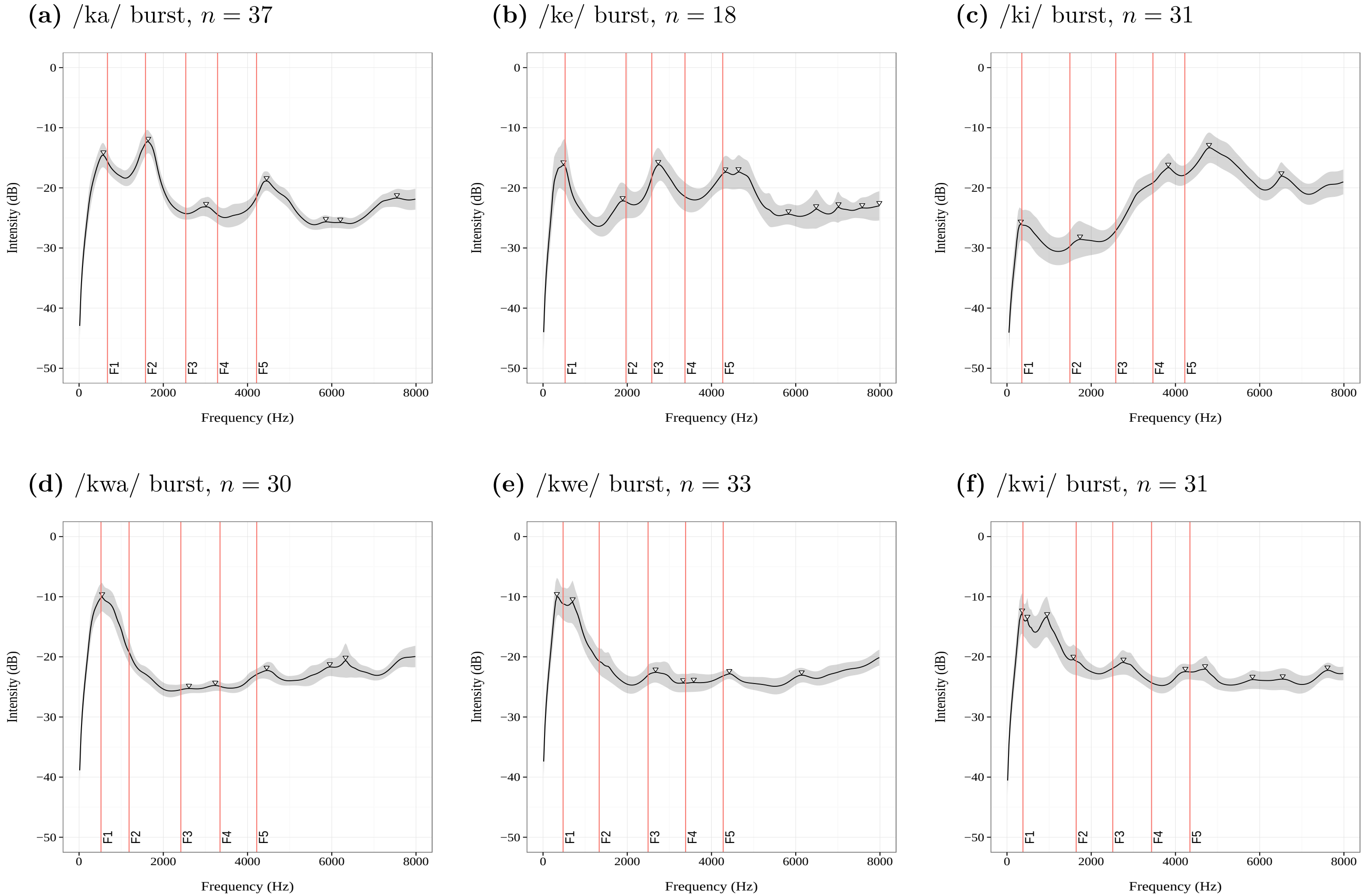


Fig. 3.16: Formant frequency transitions for the vocalic interval of Zapotec CV clusters of /kw, k/ with vowels /a, e, i/, speaker ZPP-MAT. Shaded areas indicate standard error.

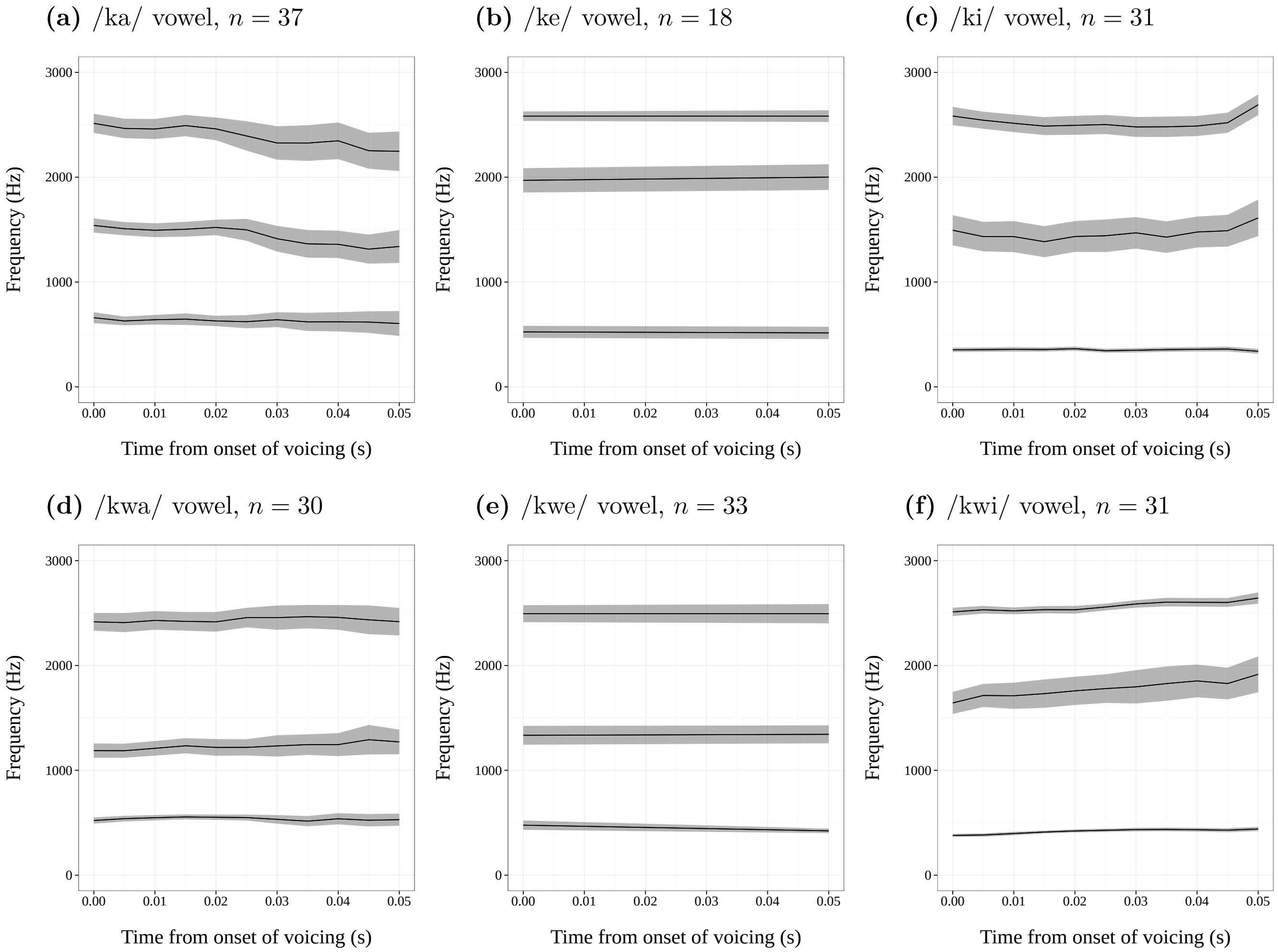
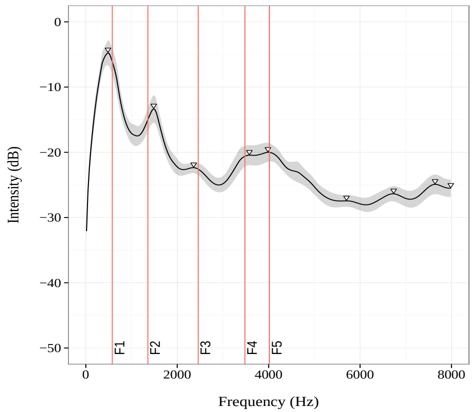
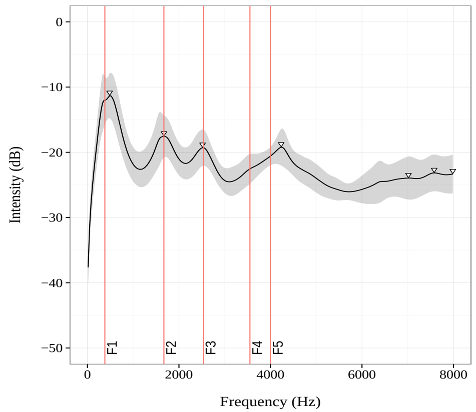


Fig. 3.17: Mean LPC spectra at the burst for Zapotec CV clusters of /kw, k/ with vowels /a, e, i/, speaker ZPP-PED. Shaded areas indicate 95% confidence interval. Arrows indicate spectral peaks. Vertical lines show mean formant frequencies measured at the point of voicing onset for each CV vocoid.

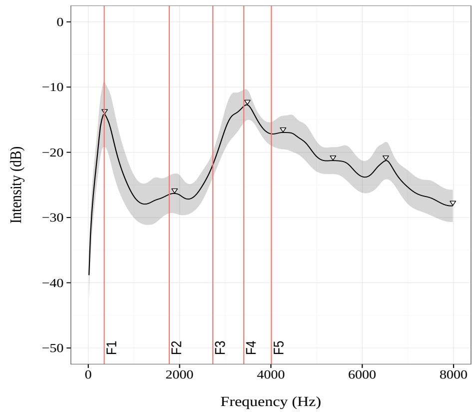
(a) /ka/ burst, $n = 28$



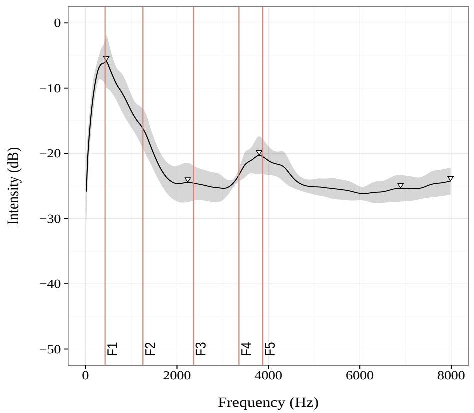
(b) /ke/ burst, $n = 12$



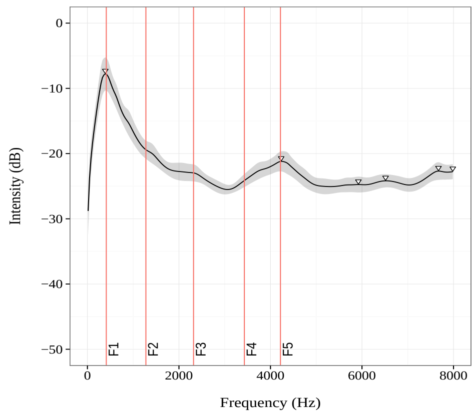
(c) /ki/ burst, $n = 12$



(d) /kwa/ burst, $n = 5$



(e) /kwe/ burst, $n = 25$



(f) /kwi/ burst, $n = 12$

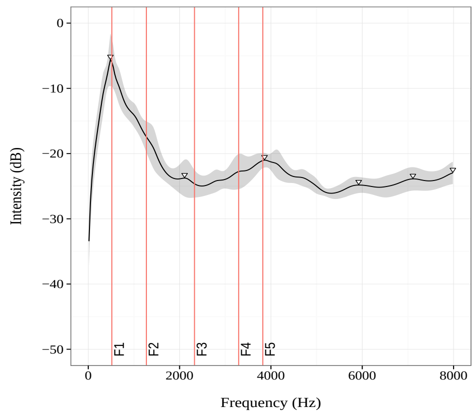


Fig. 3.18: Formant frequency transitions for the vocalic interval of Zapotec CV clusters of /kw, k/ with vowels /a, e, i/, speaker ZPP-PED. Shaded areas indicate standard error.

(a) /ka/ vowel, $n = 28$ **(b)** /ke/ vowel, $n = 12$ **(c)** /ki/ vowel, $n = 12$

(d) /kwa/ vowel, $n = 5$ **(e)** /kwe/ vowel, $n = 25$ **(f)** /kwi/ vowel, $n = 12$

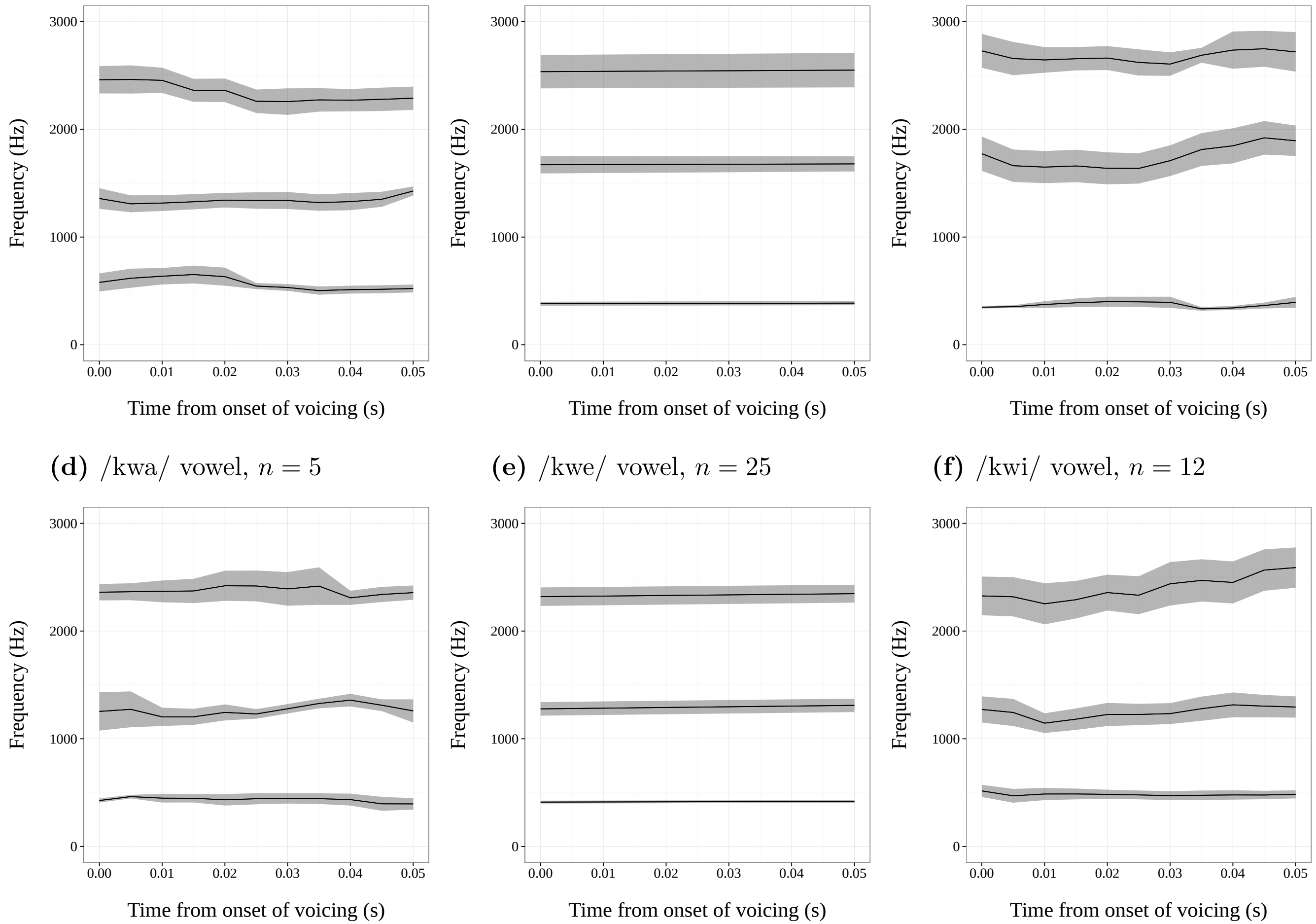
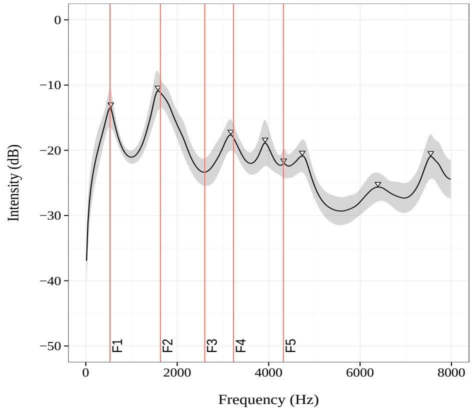
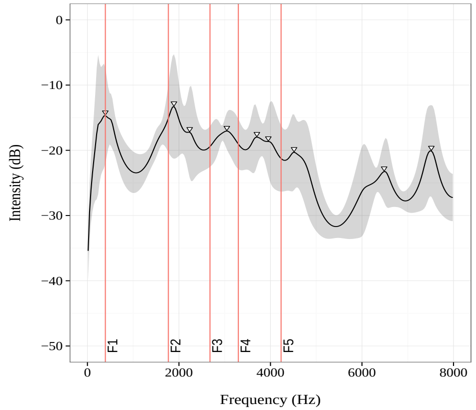


Fig. 3.19: Mean LPC spectra at the burst for Zapotec CV clusters of /kw, k/ with vowels /a, e, i/, speaker ZPP-ANT. Shaded areas indicate 95% confidence interval. Arrows indicate spectral peaks. Vertical lines show mean formant frequencies measured at the point of voicing onset for each CV vocoid.

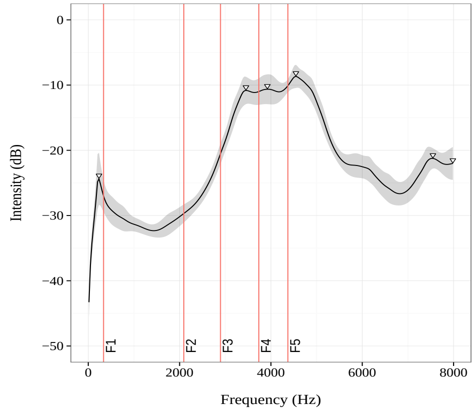
(a) /ka/ burst, $n = 15$



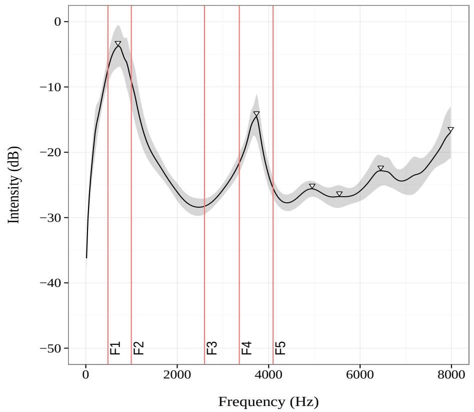
(b) /ke/ burst, $n = 5$



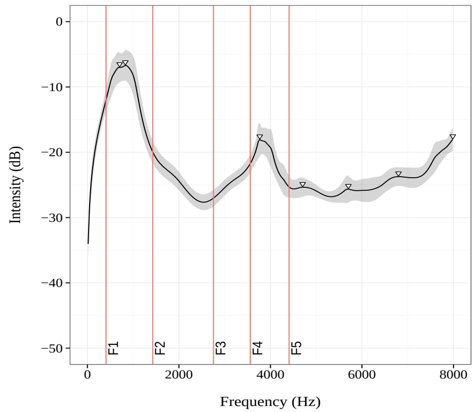
(c) /ki/ burst, $n = 17$



(d) /kwa/ burst, $n = 15$



(e) /kwe/ burst, $n = 19$



(f) /kwi/ burst, $n = 6$

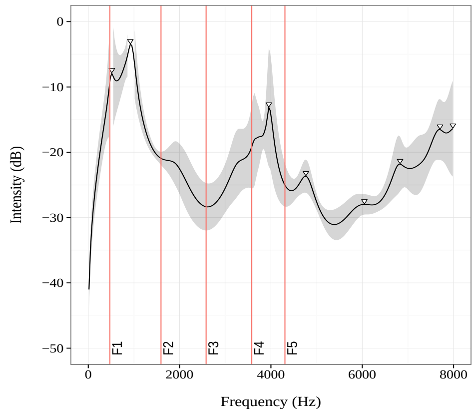
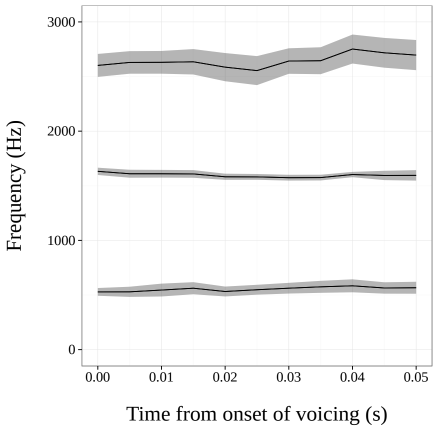
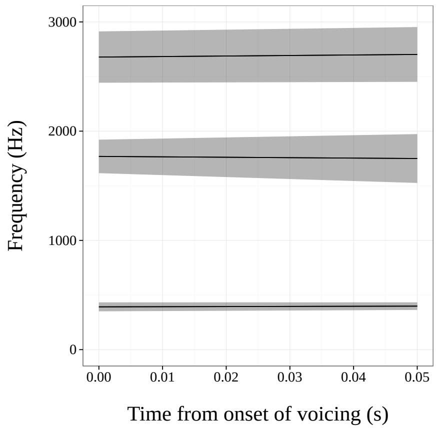


Fig. 3.20: Formant frequency transitions for the vocalic interval of Zapotec CV clusters of /kw, k/ with vowels /a, e, i/, speaker ZPP-ANT. Shaded areas indicate standard error.

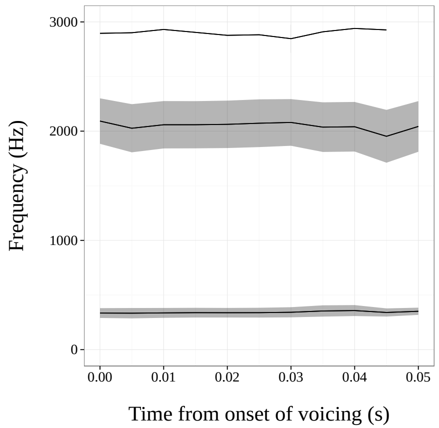
(a) /ka/ vowel, $n = 15$



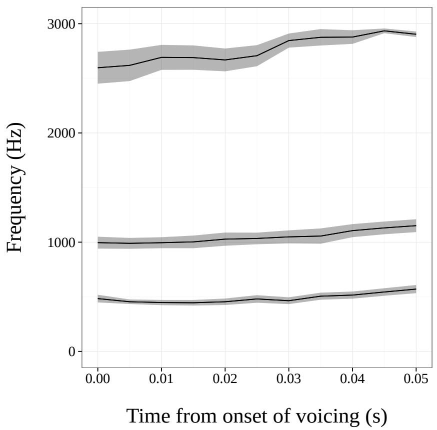
(b) /ke/ vowel, $n = 5$



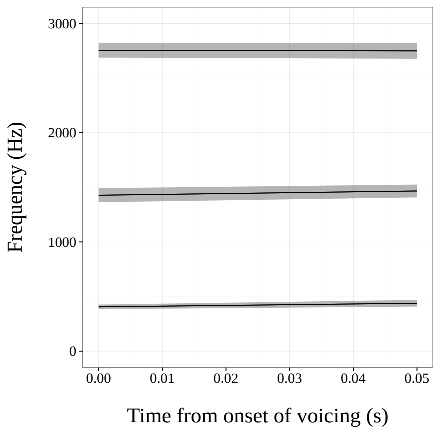
(c) /ki/ vowel, $n = 17$



(d) /kwa/ vowel, $n = 15$



(e) /kwe/ vowel, $n = 19$



(f) /kwi/ vowel, $n = 6$

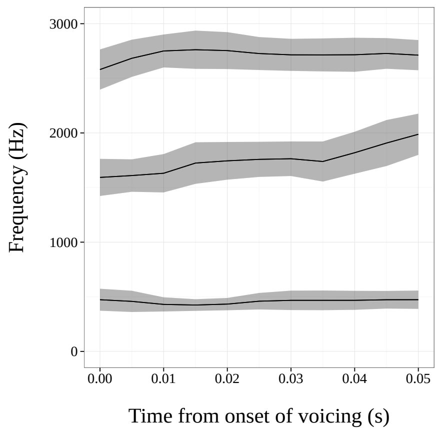


Table 3.17: Mean formant frequencies for the /V/ segment of each Zapotec /CV/ cluster in the study, measured at the onset of voicing (Hz).

	ZPP-MAT					
	n	$\bar{F}_1$	$\bar{F}_2$	$\bar{F}_3$	$\bar{F}_4$	$\bar{F}_5$
/ka/	23	674	1577	2534	3288	4214
/kwa/	26	521	1188	2416	3347	4218
/ki/	24	353	1494	2583	3465	4222
/kwi/	23	379	1643	2511	3433	4344
/ke/	10	525	1970	2582	3370	4267
/kwe/	23	477	1334	2494	3385	4278
	ZPP-PED					
	n	$\bar{F}_1$	$\bar{F}_2$	$\bar{F}_3$	$\bar{F}_4$	$\bar{F}_5$
/ka/	28	571	1341	2345	3446	4033
/kwa/	11	437	1264	2375	3434	3987
/ki/	12	372	1746	2679	3500	4102
/kwi/	12	483	1247	2399	3341	3935
/ke/	12	384	1675	2542	3554	3989
/kwe/	19	422	1447	2751	3562	4413
	ZPP-ANT					
	n	$\bar{F}_1$	$\bar{F}_2$	$\bar{F}_3$	$\bar{F}_4$	$\bar{F}_5$
/ka/	15	553	1598	2640	3366	4339
/kwa/	15	487	1047	2761	3467	4287
/ki/	17	341	2050	2907	3668	4386
/kwi/	6	457	1752	2713	3688	4382
/ke/	5	395	1759	2690	3231	4246
/kwe/	19	422	1447	2751	3562	4413

3.6.2 Discussion

LPC spectra for Zapotec /ka/ confirm the results in Keating & Lahiri (1993) in that, for velar bursts before a back vowel, spectral peaks with the highest amplitudes are found in approximately the same frequency range as $\bar{F}_2$ of the following vowel (Figs. 3.15b, 3.17a, and 3.19a). There is an additional peak at a slightly higher frequency than that of F_5, most likely resulting from frication energy.

According to the model proposed by Keating & Lahiri (1993), velars with coarticulatory fronting will have a spectral peak between F_2 and F_4 of a following /i/ vowel (p.

98). For speaker ZPP-MAT, velars before /i/ had a relatively low-intensity spectral peak in line with $\bar{F}_2$ of /i/, and a higher spectral peak at 3828 Hz, somewhat higher than the vowel $\bar{F}_2$ of 3245 Hz (3.15c and Table 3.17). The highest energy peak for /ki/ has a mean frequency of 6516 Hz, probably resulting from increased fricativization. The high frequency concentration of energy in these tokens indicates an advanced palatal tongue position during the burst.

The spectra for /k/ before /e/ also show evidence of tongue fronting. There is a prominent peak at 2734 Hz (Fig. 3.15b), close to $\bar{F}_3$ of the /e/ vowel at 2582 Hz (Table 3.17).

An effect of secondary labialization on all vowels for all Western Zapotec speakers is lowering of F_2 frequency (Fig.s 3.16, 3.18, and 3.20), with the exception of speaker ZPP-MAT for whom /i/ in fact had a higher F_2 frequency following /kw/ than following /k/. Speakers ZPP-MAT and ZPP-ANT show a rising F_2 trajectory following /kw/, similar to that seen for the British English data.

For all speakers, the burst spectra for contrastively labialized velars were homogeneous across the different vowel environments (Figs 3.16, 3.18, and 3.20). For labialized velar bursts in all the vowel environments, the spectra show a dominant low frequency peak equivalent to F_1 of the following vowel, and another less prominent peak at a frequency between F_4 and F_5 of the following vowel. This differs from the burst spectra for the non-labialized velars, which show acoustic evidence of vowel-specific coarticulation. For non-labialized velars before /a/ and /e/, the spectra are diffuse with prominent peaks at frequencies equivalent to the formant frequencies of the following vowel; for burst spectra before /i/, there is a prominent peak at F_1 of the following /i/ followed by a trough, and then energy concentrated in the region of 3000-5000 Hz.

3.7 Conclusions

Acoustic comparisons In Experiment 1, corpus data from the Audio British National Corpus was used to compare the acoustics of British English stops with velar, labial, labialized velar, and alveolar articulations in onset positions preceding back rounded vowels and front unrounded vowels. Labialized velars have been characterized cross-linguistically as having a low second formant frequency, but it was found that third formant frequency is a more prominent feature of differentiation in back rounded vowel contexts. Plain velars with non-distinctive labialization resulting from proximity to a following back rounded vowel did not differ significantly from labialized velars in formant transition frequencies. In Experiment 2, data from speakers of British English recorded in the laboratory allowed inter-speaker comparison and confirmed that labialization causes lowering of F_2 frequencies. Additionally, it was shown that voice onset time was longer for stops with either coarticulatory or contrastive labialization, but that voice onset time may not be a salient cue to place. The degree of acoustic similarity between /kwo/ and /ko/ was found to differ between speakers.

In both experiments, front vowels following /kw/ had lower formant frequencies than front vowels in positions following the other stops tested. This effect was stronger for vowels with more fronting (i.e. a higher F_2 frequency, as can be seen from the dynamic comparison of /ɪ/ versus /i/ after /kw/ in Experiment 1). There were fewer context-specific vowel differences amongst back vowels; this is presumably because there was lip-rounding during all stop releases for the back vowel tokens, rather than only being present for /kw/ articulations as in the front vowel environments.

Regarding the acoustic similarities between /kwo/ and /po/ which have been suggested as a basis for perceptually motivated sound changes, there seems to be little evidence to support the proposition that obvious or important acoustic similarities could have contributed to these sound changes. Back vowels had generally lower formant frequencies after /kw/ than after /p/; this was consistent in both experiments. For the BNC data, RMS amplitudes of /p/ were more similar to /k/ than to /kw/ in all vocalic environments. Experiment 2 found that for some speakers /kwo/ is more acoustically similar to /to/ than to /po/ (Fig. 3.13). However, it is possible that there is a language-specific element at play; for example, if ancient Greek had a more palatalized realization of /to/ (that might have been more different from /po/ than in English), F_2 on release would be higher than for English /to/, resulting in a greater difference between /kwo/ and /to/ and a relatively greater similarity between /kwo/ and /po/.

In Experiment 3, burst spectra and formant transitions for contrastively labialized and non-labialized velars preceding vowels /a, e, i/ in Western Zapotec were compared across three speakers. Similarities with British English in formant transitions for /i/ after /kw/ were observed; for two speakers, /i/ had a rising F_2 trajectory after /kw/. Whilst this was not observed for one of the speakers, the high number of speakers in the British English data could conceal similar speaker-to-speaker differences whilst revealing a more general tendency towards dynamically rising F_2 frequency for /kwi/ sequences. Likewise, for all speakers and CV sequences—with the exception of ZPP-MAT's /ki/ and /kwi/ tokens—vowels following a labialized velar had lower F_2 frequencies than vowels following a non-labialized velar, as in the British English examples. This is indicative of coarticulatory lip-rounding extending through the vowel, increasing the length of the vocal cavity and resulting in lower F_2 frequency. More striking is the difference in context-dependent ranges of variation in the bursts; for Western Zapotec speakers, there was less context-dependent variation in labialized velar burst spectra across the different vowel environments than for British English speakers. More specifically, there was some indication of coarticulatory fronting of the burst for contrastively labialized velars before /i/ in British English, but not for Zapotec.

The results suggest that there is probably not a greater likelihood of perceptual confusion of /kwo/ with /po/ than for other stops in this vocalic position. The significant lowering of F_3 frequency after /kw/ found in the BNC data, and the significant lowering of F_2 frequency after /kw/ found in both datasets, indicate acoustic dissimilarity between /kw/ and /p/ before /o/. Likewise, /kwi/ and /ti/ were dissimilar at voicing onset in Experiments 1 and 2, and the formant transition analysis in Experiment 1 found that /kw/ and /t/ are dynamically distinct across the duration of the vocoid /i/. This indicates that perceptual confusion of /kwi/ and /ti/ is highly unlikely, making the acoustic misperception account of Greek changes of /kw/ to /t/ before front vowels implausible.

The acoustics of /kw/ and /k/ in back round vowel contexts provide a strong indication of potential for perceptual confusion that could lead to the typologically common development of delabialization of labialized velars in back round vowel environments. In both experiments, labialized velars and velars showed statistically indistinguishable formant frequencies at the onset of voicing; Experiment 1 showed that these similarities continued dynamically over the entire vocoid. This acoustic similarity was not observed for non-round back vowel environments; in the Zapotec data, burst spectra showed very different energy distributions for labialized and non-labialized velars in non-round back vowel environments.

Velar fronting For plain velars, burst spectra were comparable across both languages; before /i/, the /k/ bursts of both Zapotec and English were characterized by a prominent peak equivalent to F_1 of the following vowel, and subsequent prominent peaks at frequencies above that of F_3 of the following vowel. The prominence of the high frequency peaks was greater for the Zapotec tokens than for the British English tokens. This confirms the results in Keating & Lahiri (1993) that fronted velar bursts are characterised by dominant peaks between F_2 and F_4 of a following front vowel (§3.2.1). It was predicted that labialized velars would show less evidence of velar fronting than non-labialized velars before front vowels owing to the need to retract the tongue during the articulation of secondary labiovelarization (§3.2.7) and the burst spectra presented here partly confirmed that prediction; non-labialized velars showed a high energy peak between F_2 and F_4 of a following /ɪ/ for British English and /i/ for Western Zapotec, whilst labialized velars had their most prominent peak at a frequency equivalent to F_1 of the following vowel in both languages. However, before a high front vowel /i/ in British English, the labialized velar burst does in fact have prominent peaks between F_2 and F_4 of a following vowel, and a dominant peak between F_4 and F_5 (Fig. 3.7c). This constitutes acoustic evidence of fronting of velars with contrastive secondary labialization in front vowel environments, and will be discussed further with reference to historical sound change in §5.3.

Gesture timing As discussed in §3.2.3, languages with contrasts between sequences of a stop and an approximant, and stops with secondary articulations (e.g. /pj/ vs /pʲ/) appear to use differences in gesture timing as cues to differentiate these sounds. The languages analysed here differ in how the effects of secondary labialization are manifested in the acoustic signal: for British English, there is evidence of coarticulatory velar fronting in the burst, but highly dynamic formant transitions indicating lowering of F_2 during the onset of the vowel under the effects of lip-rounding. For Western Zapotec, the situation is reversed; lip-rounding amplifies the low frequency range of the spectrum, but does not have a strong lowering effect on following vowel formants. This suggests that velar articulation and lip-rounding are more sequential for British English and more simultaneous for Western Zapotec. The observation makes the British English data particularly interesting from the point of view that the front vowel exerts an anticipatory coarticulatory effect on the velar despite the intervening glide.

Predictions The acoustic analysis shows that labialized velars are acoustically distinctive in front vowel environments; in both British English and Western Zapotec there is a lowering effecting on the second format frequency of a following front vowel, although this is stronger for English than for Zapotec. At the release of the velar closure, there is evidence of relative fronting of the velar place of articulation for British English, demonstrated by a high frequency peak in the burst spectrum. This effect is less obvious for Western Zapotec, in which the overlap of the velar burst and secondary labialization causes a dominant low frequency peak in the burst spectra of two speakers, whilst the third speaker does have a high frequency peak similar to that observed for /ki/. With reference to the Greek sound changes of /kwi/ to /ti/, it is not clear whether the fronting at the burst could trigger misperception of /kwi/ as /ti/ or /ki/; it seems likely that the continued presence of low frequency energy in the burst, coupled with the dynamic rising of F_2 frequency through the vocoid, may provide sufficient cues to enable the listener to identify the labialized velar place of articulation.

Labialized velars and labials are more acoustically similar to one another in back round vowel environments than in front vowel environments. However, labialized velars and labials were not more similar to one another than to other stops in back round vowel environments. Contrastively labialized and non-contrastively labialized velars had similar formant transitions, but differed in their burst spectra.

In the next chapter, a perceptual confusion experiment investigates how the acoustic differences and similarities exposed in Experiments 1 and 2 are realised perceptually. If acoustic similarity does correspond to perceptual similarity, the data presented in this chapter suggest that listeners will not display a perceptual bias towards confusing /kwo/ with /po/ and /kwi/ with /ti/, but may be biased towards confusing /kwo/ and /ko/.

4

Perception of Labialized Velars

4.1 Experiment 4: Perceptual confusion analysis

In the previous chapter, obstruents in British English and Western Zapotec were analyzed in order to investigate the acoustic similarity between labialized velars and labials, coronals, and non-labialized velars. It has been claimed that acoustic similarity between labialized velars and labials, and between contrastively labialized velars and non-labialized velars in labializing environments, might cause sound changes of labialized velars to labials and velars by provoking misperception during speaker-listener interactions. Consequently, in this chapter I use the results of the acoustic study to inform a perceptual study which investigates how the acoustic characteristics described in Chapter 3 interact with perception. The experiment does this by exposing participants to stimuli with the obstruents /p, t, k, kw/ and vowels /e, i, o/ and inducing misperception in order to identify perceptual biases (cf. similar methodologies in, amongst others, Miller & Nicely 1955 and Winitz et al. 1972).

Experiment 1 found that there was no particular acoustic similarity between labialized velars and labials, indicating that there is not likely to be perceptual similarity between these sounds. This finding prompts the following hypothesis to be proposed for the perception study:

Prediction 1: Participants are not more likely to confuse labialized velars with labials than with other stops.

Although a perceptual confusion account has not been proposed for sound changes of labialized velars to coronals, these sounds are included in the perceptual study in order to ensure the investigation is not affected by subjective presuppositions. The acoustic study indicated that labialized velars before front vowels can show acoustic evidence of coarticulatory fronting which manifests as a peak above 2.5 kHz on the burst spectrum, but that in most instances (including before the British English close-mid front vowel /ɪ/) this is overshadowed by dominant low frequency peaks. In addition to this, the formant transitions of British English front vowels after labialized velars are highly dynamically distinctive, with a steeply rising F_2 transition. Thus it is not expected that participants will mistake labialized velars for coronals:

Prediction 2: Participants are not more likely to confuse labialized velars with coronals than with other stops before front vowels.

Contrastively labialized velars were highly acoustically similar to non-contrastively labialized velars before back round vowels. Acoustic cues to contrastive labialization were mostly encoded in the burst rather than in formant transitions, but were still not highly distinctive. For this reason it is expected that there will be a high degree of perceptual confusion between contrastively and non-contrastively labialized velars in positions preceding back round vowels.

Prediction 3: Participants are more likely to confuse contrastively labialized velars with non-contrastively labialized velars than with other stops before back round vowels.

The investigation was limited to British English because it was not possible to conduct a parallel experiment with Western Zapotec speakers due to time constraints. Justification for the use of British English as an empirical proxy to inform conclusions regarding Ancient Greek and other languages can be found in §1.5.

4.2 Methodology

4.2.1 Stimulus material

Stimulus words were selected as described in §3.5.

Distracter words, which were minimally contrasting but did not include the stops selected for discussion (e.g. *sod*, *lip*), were included in order to prevent participants easily guessing the focus of the experiment and to alleviate boredom (Table 4.1).

Table 4.1: Distracter words used in Experiment 1.

bail	bill	bit	born	bought	dale	dawn	dill
dip	dressed	fail	fake	fought	fawn	fill	fit
fraught	jest	lake	lest	lick	lip	mail	make
male	mill	mod	mourn	nest	nip	nod	prawn
prick	prod	rest	sail	sake	sawn	short	sick
sill	sip	sit	sod	sort	spit	trill	twit
veil	wad	wake	warn	wart	west	whale	whip
wick	will	wit	zest	zip			

Participants were presented with 360 stimuli, of which 150 were target words and 210 were distracters.

Unfortunately, due to a paucity of appropriate English minimal pairs, it was not possible to give participants the opportunity to compare words in the /ke/ and

/pe/ groupings with one another. For example, there is no English word *[pʰɛɪk] to correspond with *cake*, *quake*, and *take*, nor is there a *[kʰɛst] corresponding to *quest*, *pest*, and *test*. With this exception, subjects had multiple opportunities to compare every CV grouping with every other CV grouping numerous times.

Table 4.2: Response options available to participants for each CV stimulus grouping. It was not possible to compare /ke/ and /pe/ words due to a lack of appropriate English minimal pairs.

			Response											
				p			t			k			kw	
			pe	pi	po	te	ti	to	ke	ki	ko	kwe	kwi	kwo
Stimulus	p	pe				12			—			12		
		pi					6			12			6	
		po						6			12			6
	t	te	6						6			6		
		ti		9						6			9	
		to			6						12			6
	k	ke	—			12						6		
		ki		9			6						9	
		ko			12			6						
	kw	kwe	6			12			6					
		kwi		9			6			9				
		kwo			12			12			12			

4.2.2 Preparing stimuli

The spoken words were recorded on one channel at 44.1 kHz (CD quality) but a sampling rate of 16 kHz is adequate for speech recording, so the recordings were down-sampled and the empty channel removed. The long recordings were manually edited into single-word files. Several different methods of altering or obscuring the signal to provoke misperception were considered, including selective band-stop filtering to obscure specific formant information; this was rejected due to the difficulty of accounting for characteristic differences in speakers' individual formant patterns and the danger of 'pushing' participants in a particular direction by rendering certain sounds too similar to one another and thus introducing bias to the experiment (see Miller & Nicely 1955 for a discussion of the effects of low- and high-pass filtering on perception of English stops). Obscuring the signal with natural, conversational 'café' noise was selected as the most appropriate method of proceeding (after Heinrich et al. 2010; see Bronkhorst 2000 and Heinrich et al. 2008 on the effectiveness of conversational babble in masking the speech signal and distorting perception). This approach meant that misperception was prompted without tampering with the underlying speech signal, in contrast with e.g. band filtering which could disproportionately affect some stimuli, depending upon the frequency parameters of the filter. The random distribution of different amplitudes and frequencies inherent to the conversational babble reflects the type of obscuring noise encountered naturally by speakers; this was thought to be an advantage in the experiment which would not have been achieved by the introduction of a regularly repeating artificial buzz or hum. White noise was rejected because the range and distribution of amplitudes within the conversational noise was thought to be necessarily equivalent to those found within the stimulus material, given that both are representative of actual human speech; had artificially produced noise been used to mask the stimulus material, this would not have been the case. Additionally, the production of artificial noise would have required the selection of specific frequency and energy parameters which could result in unintentional bias, or in noise which masked the different stimulus sounds unequally.

Recorded conversational babble was sourced from the internet[1] and the sampling rate altered to match that of the laboratory recorded speech (i.e. down-sampled to 16 kHz). The available clips were not longer than 90 seconds long, so three were concatenated to produce one long audio file. In order to control for the naturally inconsistent nature of the conversational noise (some parts were naturally quieter or louder than others) eight copies of each recorded word were made; for each of these a segment of corresponding duration was taken from the conversational noise, with start points selected at random. This strategy was chosen because it was thought that the best way to control for this variable — the naturally erratic

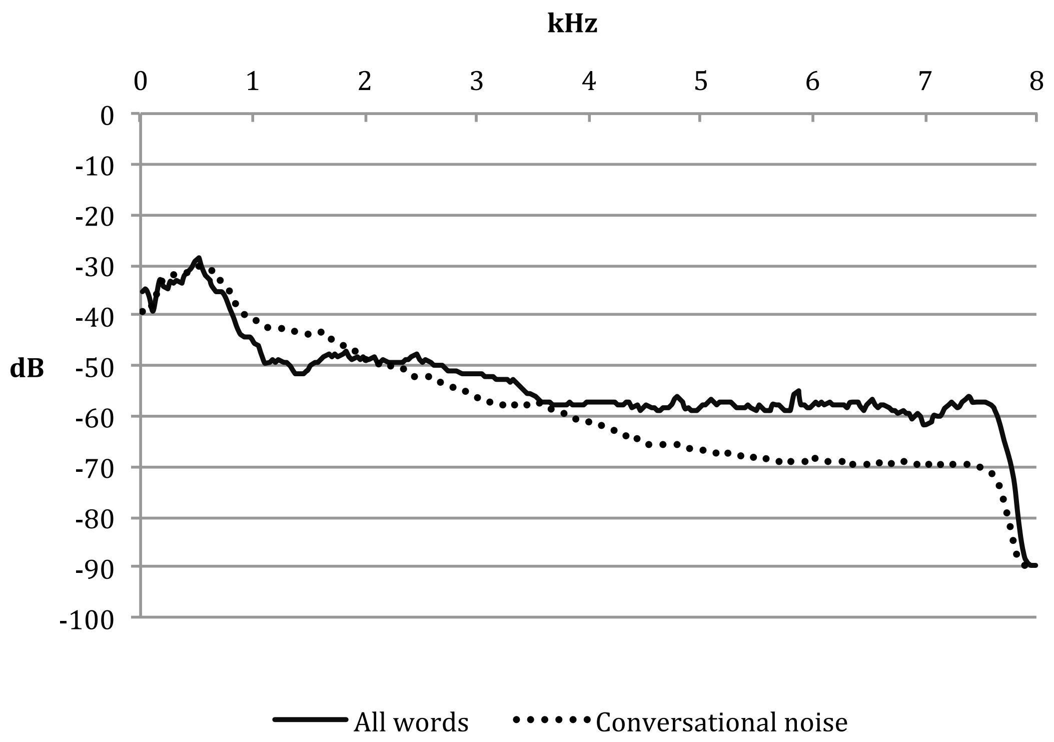
[1] http://www.soundjay.com/human/crowd-talking-1.wav; http://www.soundjay.com/human/crowd-talking-2.wav; http://www.soundjay.com/human/crowd-talking-3.wav.

energy patterns within the background noise — was to randomize it and thus expose participants to background noise of randomly varying frequencies within the range of human speech, more accurately representing real-life conditions for misperception.

The volumes of the combined word and noise files were then normalized to 50% of the maximum amplitude in order to control for volume differences between individual speakers and words, and between the words and noise. The recorded words were then mixed with their corresponding noise samples, giving a combined maximum amplitude of 100% of the 16-bit range (0 dB). This yielded a pool of 48 unique combined noise and audio files for each recorded word (given that for each word recorded by each of six speakers, eight unique noise files were created and combined with the words). Different signal-to-noise ratios were informally tested during initial practice runs of the experiment; a ratio of 1:1 resulted in approximately 80% correct responses, leaving relatively little incorrect response data with which to work. A ratio of 3:7 resulted in a high proportion of incorrect responses but left participants demoralized and less willing to concentrate on the task. For these reasons it was determined that a 2:3 ratio (with an intensity level of –1.76 dB), yielding an approximate 30% incorrect responses, was appropriate for this pilot.

A comparison of the long term average spectra for the stimulus material and the masking noise is shown in Figure 4.1. The intensity of the conversational noise equalled or exceeded that of the first to third formants of the stimulus words. The intensity of the stimulus words exceeded that of the masking noise in the region 4–6 kHz; the high relative intensity of the stimuli could be due to the fact that all stimuli consisted of voiceless stops with high frequency aspiration, raising the average intensity of the stimuli slightly above that of normal speech but reflecting the 'naturalness' of the masking noise, in that it shows the intensity of the stimuli relative to other speech.

Fig. 4.1: Long term average spectra for the stimulus material and the conversational noise used for masking. The conversational noise masks the first to third formants of the stimulus material.



4.2.3 Experiment structure

The experiment was conducted using E-Prime software. Participants were seated in front of a computer screen in a soundproofed recording booth at the Phonetics Laboratory, and heard the stimulus material through headphones. In each trial, three single-word options were presented on the screen, with each option corresponding to a number which participants were asked to select using the number keys on a normal computer keyboard. Following a delay of one second after the option words appeared on the screen, the stimulus audio was played; stimulus audio for each word was randomly selected from the pool of 48 unique files described in §4.2.2. Participants were informed that one of the options was always the correct answer. Due to the fact that three options were presented, but there were four potential stops, options were rotated equally to counteract potential bias resulting from the selection environment. For example, when *pe* was the input, there were an equal number of trials in which the two incorrect answers were *te, ke*, or *te, kwe*, or *ke, kwe*; note that all response option opportunities in 4.2 are multiples of three.

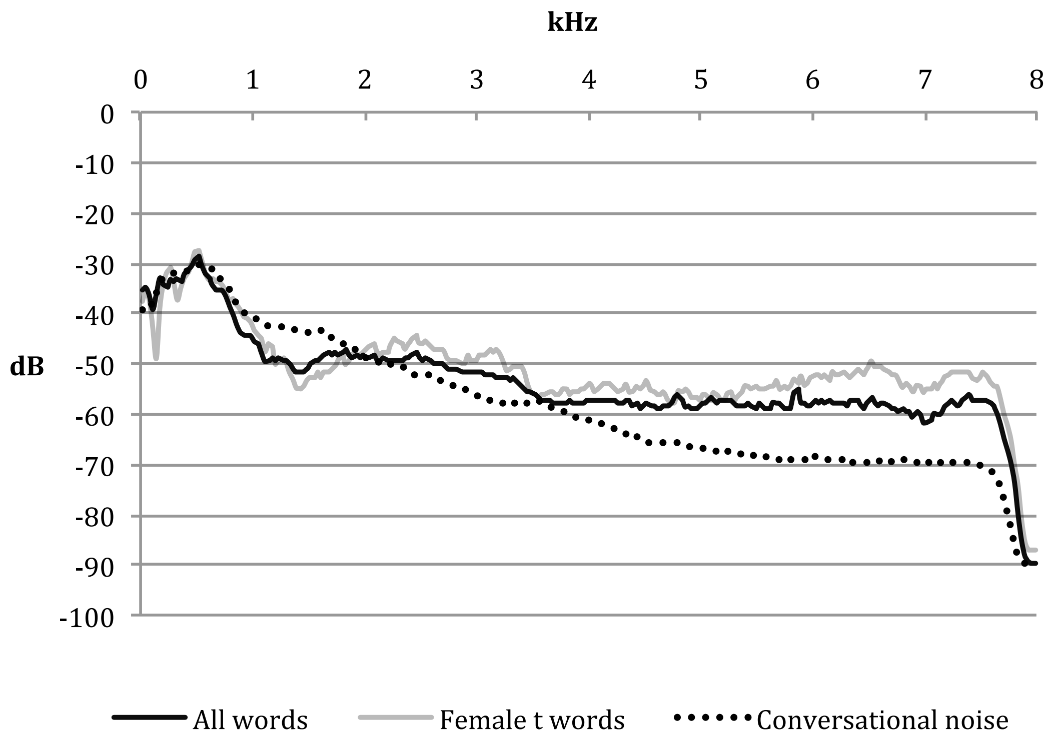
The options remained on the screen until the participant made a selection, after which the experiment progressed immediately to the next trial. Trials were presented in a randomized order unique to each participant in order to control for experiment fatigue, which might otherwise have caused data gathered near the end of the experiment to be less reliable than that at the beginning.

Twelve participants were recruited from the local population in Oxford and screened for dialect consistency (standard Southern British English) to ensure the dialect background of the stimulus speakers and the subjects was as similar as possible. The ages of the participants ranged from 19 to 69, with a mean age of 32.6 years; six were male and six female. Subjects took approximately 35 minutes to complete the experiment and were reimbursed for their participation.

4.3 Results

The primary difference in perceptual confusion patterns between individual speakers was found in responses to stimuli in the /te/, /ti/ and /to/ groups, in which responses to male speakers showed some confusion but responses to female speakers were overwhelmingly accurate. This could be an effect of the conversational noise used to mask the stimuli; when the long term average spectra for the background noise and for the male and female tV stimuli are compared to one another, there is a high energy peak in the range 2.5–3.5 kHz for the female speakers which is not masked by any equivalent peaks in the background noise (see Figure 4.2). Although this is in line with previous perceptual confusion experiments in English—which also found high levels of accuracy in [t] identification (Miller & Nicely, 1955; Winitz et al., 1972)—due to the disparity in identification rates for male and female speakers for words in the tV groups and the potential role of the masking noise, and the typically significant differences in acoustic measurements for males and females shown in Chapter 3, responses to males and females will be treated separately.

Fig. 4.2: Long term average spectra for all stimulus words against background noise and tV group words elicited from female speakers, demonstrating how high frequency peaks in the 2.5–3.5 kHz range could have made female tV group words easier to perceive.



The response data for each word were compiled into a confusion matrix consisting of absolute numbers of misperceptions for each stimulus/response combination. As explained in §3.5.2, each word was assigned to a CV group to reflect the conditioning environments for misperception. Further confusion matrices were derived from the experimental data by dividing the total instances of a CV_r response to a CV_s stimulus by the total number of trials in which CV_s was the stimulus and CV_r a possible response.

I.e.

$$\text{Confusion rate} = \frac{\text{Number of y responses to x stimuli}}{\text{Number of trials where x is the stimulus and y is an option}} \tag{4.1}$$

Results of this analysis, broken down by individual speaker, are presented in Table 4.3.

Table 4.3: Identification probabilities: subject responses, broken down by individual speaker. Stimuli are organised by *CV* grouping, as a proportion of all responses to the stimulus *CV* where the response *CV* was an available option.

		Response											
			p			t			k			kw	
		pe	pi	po	te	ti	to	ke	ki	ko	kwe	kwi	kwo
Speaker 1F	kwe	0.00			0.00			0.11			0.95		
	kwi		0.07			0.00			0.19			0.78	
	kwo			0.00			0.20			0.38			0.52
Speaker 2F	kwe	0.00			0.00			0.00			1.00		
	kwi		0.06			0.00			0.00			0.96	
	kwo			0.00			0.25			0.00			0.82
Speaker 3F	kwe	0.00			0.00			0.07			0.96		
	kwi		0.21			0.00			0.05			0.79	
	kwo			0.14			0.06			0.41			0.50
Speaker M1	kwe	0.00			0.00			0.20			0.92		
	kwi		0.00			0.00			0.00			1.00	
	kwo			0.21			0.17			0.23			0.58
Speaker M2	kwe	0.14			0.08			0.42			0.65		
	kwi		0.10			0.00			0.15			0.80	
	kwo			0.00			0.10			0.14			0.81
Speaker M3	kwe	0.08			0.07			0.00			0.90		
	kwi		0.13			0.00			0.09			0.85	
	kwo			0.16			0.07			0.27			0.61

Due to the importance of listener-specific perceptual mapping in speech perception (Beddor, 2012; Solé, 2014) response ratios are broken down by individual listening participant in Table 4.4.

Table 4.4: Confusion matrices for responses to *kw* stimuli broken down by listener.

		Response											
			p			t			k			kw	
		pe	pi	po	te	ti	to	ke	ki	ko	kwe	kwi	kwo

Listener 1	kwe	0.00			0.00			0.33			0.83		
	kwi		0.22			0.00			0.12			0.75	
	kwo			0.00			0.00			0.50			0.50
Listener 2	kwe	0.00			0.00			0.17			0.92		
	kwi		0.00			0.00			0.00			1.00	
	kwo			0.00			0.17			0.17			0.75
Listener 3	kwe	0.00			0.00			0.00			1.00		
	kwi		0.11			0.00			0.00			0.92	
	kwo			0.00			0.00			0.33			0.67
Listener 4	kwe	0.17			0.00			0.17			0.83		
	kwi		0.00			0.00			0.00			1.00	
	kwo			0.00			0.00			0.50			0.50
Listener 5	kwe	0.00			0.00			0.00			1.00		
	kwi		0.11			0.00			0.00			0.92	
	kwo			0.33			0.33			0.00			0.67
Listener 6	kwe	0.17			0.00			0.00			0.92		
	kwi		0.11			0.00			0.00			0.92	
	kwo			0.00			0.17			0.25			0.67
Listener 7	kwe	0.17			0.00			0.17			0.83		
	kwi		0.33			0.00			0.33			0.50	
	kwo			0.33			0.00			0.42			0.42
Listener 8	kwe	0.00			0.17			0.33			0.67		
	kwi		0.00			0.00			0.00			1.00	
	kwo			0.00			0.17			0.42			0.50
Listener 9	kwe	0.00			0.00			0.01			1.00		
	kwi		0.11			0.00			0.00			0.92	
	kwo			0.33			0.17			0.17			0.58
Listener 10	kwe	0.00			0.00			0.17			0.92		
	kwi		0.22			0.00			0.22			0.67	
	kwo			0.17			0.00			0.42			0.50
Listener 11	kwe	0.00			0.00			0.17			0.92		
	kwi		0.00			0.00			0.22			0.83	
	kwo			0.17			0.00			0.08			0.83
Listener 12	kwe	0.00			0.17			0.00			0.83		
	kwi		0.00			0.00			0.00			1.00	
	kwo			0.00			0.50			0.00			0.75

The perceptual distance between each pair of CV groupings was calculated using the method described in Shepard (1972) and Johnson et al. (2011). First, perceptual similarity was calculated by expressing the sum of the confusions as a proportion of the sum of correct responses. This approach removes asymmetrical perceptual biases.

Table 4.5: Identification probabilities based on responses to female speakers only; subject responses to stimuli by CV grouping, as a proportion of all responses to CV_s in trials where CV_r was an available option; a blank cell designates a selection that was never available to participants. As explained in §3.5.2, pe—ke was the only possibility not captured in this pilot.

			Responses to Female Speakers											
				p			t			k			kw	
			pe	pi	po	te	ti	to	ke	ki	ko	kwe	kwi	kwo
Stimulus	p	pe	0.84			0.16			—			0.00		
		pi		0.50			0.04			0.08			0.06	
		po			0.31			0.08			0.44			0.36
	t	te	0.00			0.99			0.00			0.01		
		ti		0.00			0.96			0.09			0.00	
		to			0.13			0.83			0.06			0.09
	k	ke	—			0.11			0.89			0.00		
		ki		0.06			0.19			0.79			0.09	
		ko			0.27			0.09			0.51			0.36
	kw	kwe	0.00			0.00			0.06			0.97		
		kwi		0.12			0.00			0.07			0.85	
		kwo			0.07			0.14			0.33			0.56

Table 4.6: Identification probabilities based on responses to male speakers only; subject responses to stimuli by CV grouping, as a proportion of all responses to CV_s in trials where CV_r was an available option; a blank cell designates a selection that was never available to participants. As explained in §3.5.2, pe—ke was the only possibility not captured in this pilot.

			Responses to Male Speakers											
				p			t			k			kw	
			pe	pi	po	te	ti	to	ke	ki	ko	kwe	kwi	kwo
Stimulus	p	pe	0.62			0.22			—			0.16		
		pi		0.52			0.19			0.34			0.11	
		po			0.25			0.11			0.53			0.36
	t	te	0.27			0.57			0.47			0.07		
		ti		0.24			0.60			0.27			0.11	
		to			0.17			0.38			0.37			0.32
	k	ke	—			0.06			0.85			0.09		
		ki		0.12			0.00			0.89			0.02	
		ko			0.22			0.15			0.50			0.38
	kw	kwe	0.08			0.05			0.18			0.82		
		kwi		0.08			0.00			0.08			0.89	
		kwo			0.14			0.11			0.23			0.65

$$S_{ij} = \frac{p_{ij} + p_{ji}}{p_i + p_{jj}} \tag{4.2}$$

For example, the perceptual similarity between /pi/ and /kwi/ words uttered by male speakers (data in Table 4.6) in this experiment would be calculated as follows:

$$S_{pkw} = 0.14 = \frac{0.11 + 0.08}{0.52 + 0.89} \tag{4.3}$$

The perceptual distance between the two sounds can then be calculated as the negative of the natural log of the similarity:

$$d_{ij} = -ln(S_{ij}) \tag{4.4}$$

The results of this for the aggregated speaker data are presented in Table 4.7 and Table 4.8.

The mean sensitivity index or d' of all words within each CV grouping was calculated using equation 4.5, in which Z is the inverse of the cumulative Gaussian distribution.

$$d' = Z(\text{hit rate}) - Z(\text{false alarm rate}) \tag{4.5}$$

The hit rate was calculated by dividing the number of correct identifications of a stimulus word by the number of times that word appeared; the false alarm rate is the number of times participants selected that word as a response when it was not the stimulus, as a proportion of all trials in which the word was a response option available to participants. *d'* data are presented in Table 4.9; the higher the *d'*, the more easily was the stimulus detected by the subjects.

It was discovered that words in the /te/, /ti/ and /to/ groups were considerably more identifiable when produced by female speakers, potentially resulting in an

Table 4.7: Showing perceptual distances between the CV combinations for female speakers in Experiment 1. Note that pe—ke is the only possibility not captured in this pilot. 'x' designates a combination with a confusion rate of 0.

			Perceptual Distances for Female Speakers											
				p			t			k			kw	
			pe	pi	po	te	ti	to	ke	ki	ko	kwe	kwi	kwo
Stimulus	p	pe	0.00			2.45			—			x		
		pi		0.00			3.49			0.93			2.01	
		po			0.00			1.70			0.15			0.70
	t	te				0.00			2.80			4.90		
		ti					0.00			1.85			x	
		to						0.00			2.19			1.83
	k	ke							0.00			3.42		
		ki								0.00			2.35	
		ko									0.00			0.45
	kw	kwe										0.00		
		kwi											0.00	
		kwo												0.00

artificially high *d'* value. If the values for these groups are removed from the *d'* tables provided in Table 4.9, the relative *d'* values of the different groups for male and female speakers are almost identical (Table 4.10).

4.4 Discussion

Confusion patterns relating to single articulation stops generally confirmed those found in previous perceptual confusion investigations: /pi/ to /ki/ confusion at a higher rate than /pi/ to /ti/ confusion (Winitz et al., 1972; Miller & Nicely, 1955); a high rate of accuracy in *t* identification, but a tendency to select *k* responses rather than *p* when *t* was misperceived (Winitz et al., 1972). On difference is that Winitz et al. (1972) found that *k* words were the stops most likely to be misperceived, whereas here it was found that *p* words elicited the highest rates

Table 4.8: Showing perceptual distances between the *CV* combinations for male speakers in Experiment 1. Note that pe—ke is the only possibility not captured in this pilot. 'x' designates a combination with a confusion rate of 0.

			Perceptual Distances for Male Speakers											
				p			t			k			kw	
			pe	pi	po	te	ti	to	ke	ki	ko	kwe	kwi	kwo
Stimulus	p	pe	0.00			0.90			—			1.81		
		pi		0.00			0.96			1.11			2.02	
		po			0.00			0.83			0.01			0.60
	t	te				0.00			0.98			2.48		
		ti					0.00			1.70			2.64	
		to						0.00			0.53			0.87
	k	ke							0.00			1.81		
		ki								0.00			2.88	
		ko									0.00			0.65
	kw	kwe										0.00		
		kwi											0.00	
		kwo												0.00

Table 4.9: *d'* values for each *CV* group, in increasing order, for female and male speakers. A *d'* greater than 1 indicates discrimination is occurring, with values greater than 4 indicating that discrimination is good. For female speakers, there was no *d'* value for /pe/ because the false alarm rate was 0.

CV	*d'* Females
po	1.399
ko	1.535
kwo	1.787
pi	2.241
ki	2.426
to	3.112
kwi	3.412
ke	4.013
ti	4.046
te	4.149
kwe	4.898

CV	*d'* Males
po	1.137
ko	1.224
to	1.668
kwo	1.924
pi	1.969
te	2.090
pe	2.441
ti	2.766
kwe	2.848
ki	2.949
ke	2.949
kwi	3.477

Table 4.10: *d'* values for male and female speakers, with the /te/, /ti/ and /to/ groups removed. The /pe/ group was also removed due to the lack of a value for female speakers. Relative values are similar for both sexes.

CV	*d'* Females
po	1.399
ko	1.535
kwo	1.787
pi	2.241
ki	2.426
kwi	3.412
ke	4.013
kwe	4.898

CV	*d'* Males
po	1.137
ko	1.224
kwo	1.925
pi	1.969
kwe	2.848
ki	2.949
ke	2.949
kwi	3.477

of confusion; however, in that study, correct responses for *k* increased at a faster rate relative to other stops when participants were provided with 100 ms of the following vowel—perhaps indicating a greater reliance on coarticulatory cues for stop identification—and in the present study participants had access to transition and vowel data for all trials.

For all stimuli, response patterns differed significantly from those expected by random selection, indicating that response biases for each input group were significant in all cases (whilst controlling for increased likelihood of correct answer selection; χ^2 test, $p < 0.05$).

As can be observed, words in the /kwi/ and /kwe/ groups scored highest for perceptual distinctiveness for both male and female speakers. This evidence confirms prediction 2, indicating that perceptual similarity between the two obstruents is unlikely to have played a significant role in the sound changes $k^w > t/\_i,e$ and $k^{wh} > t^h/\_e$.

Response patterns to /kwo/, /kwe/ and /kwi/ stimuli, broken down by speaker, are given in 4.3. The relative lack of individual variation observed demonstrates the advantages of selecting tokens from multiple speakers in this experiment. Use of stimulus material from just one speaker could have resulted in a skewed outcome (e.g. use of tokens only from Speaker 1M, who elicited confusion of /kwo/ for /po/ words at a rate of 0.21, would have given an artificially inflated impression of actual confusion rates). In contrast, the use of six different speakers allows comparison of individual misperception patterns and strengthens the conclusions drawn, as the lack of confusion between /ko/ and /po/ words and between /kwi/, /kwe/, and /ti/, /te/ words is shown to be consistent across all speakers. Although the acoustic analysis did sometimes show differences across different speakers, there were found to be no significant differences in the distribution of confusion ratios across individual speakers (Kruskal-Wallis test, $p > 0.05$) indicating that the perceptual confusion biases exposed in this experiment were not skewed by characteristics of individual speakers.

Despite some individual variation, the pattern of perceptual confusion is the same for /po/ responses to /kwo/ stimuli across speakers; there is no significant likelihood of confusion, with rates of /po/ responses for all speakers considerably below those expected by chance, and therefore it can be concluded that there is no perceptual bias in this direction for any speaker. This confirms Prediction 1. For the pooled speaker data, there was no tendency towards a /po/ response to /kwo/ simulus material (0.11). Words with /kwo/ were mistaken for the corresponding /po/ words at a rate approximately the same as that with /to/ words (0.13) and substantially less than the rate of confusion with /ko/ words (0.27). In fact, words in the /kwe/, /kwi/ and /kwo/ groups were among those most accurately identified by participants, with a mean correct identification probability of 0.79, and /po/ turned out to be the response participants were least likely to select in response to a /kwo/ stimulus (Table 4.3). Confusion between /po/ and /kwo/ was in fact heavily weighted in the opposite direction: there was a strong tendency for participants to identify /po/ stimuli as /kwo/ (0.36 using all response data, pooled without reference to speaker sex or listener) but not vice versa.

The data confirm Prediction 3 in that there was a tendency to confuse /kwo/ words with /ko/ words at a significant rate. Consistently high rates of confusion were found for all speakers and for all listeners, with the exception of Speaker 2F whose /kwo/ stimuli elicited no /ko/ responses. There was no corresponding confusion between /kwi/ and /ki/ words. In fact, as can be seen in Table 4.9, /kwo/ and /kwi/ words are at opposite ends of the perceptual distance scale. This corresponds closely to the pattern outlined in §2.6 in which labiovelars undergo dissimilation near back round vowels independently in a number of languages. The tendency revealed here is sufficiently strong that it may be worth investigating whether labiovelar changes near back rounded vowels are a prerequisite for developments towards velars in other conditioning environments.

Table 4.11 compares the Euclidean distances between mean formant values with the perceptual distances from Tables 4.7 and 4.8. It indicates (as may be expected) a highly significant positive correlation between acoustic distance at vocalic onset and perceptual distances for the kwV stimuli (one-tailed t test: $r = 0.97$, $p < 0.01$). For the correlation between acoustic distance at vocalic onset and perceptual distances for all CV combinations for male speakers, $r = 0.82$, $p < 0.05$. For female speakers, the correlation was less strongly positive, although still significant ($r = 0.59$, $p < 0.01$). It was impossible to calculate perceptual distances for those word groupings that were accurately identified in 100% of trials (see Tables 4.7 and 4.8). This occurred in two groups for stimuli produced by female speakers, leaving blank values for those groups. Plots and regression lines for each speaker are presented graphically in Figure 4.3.

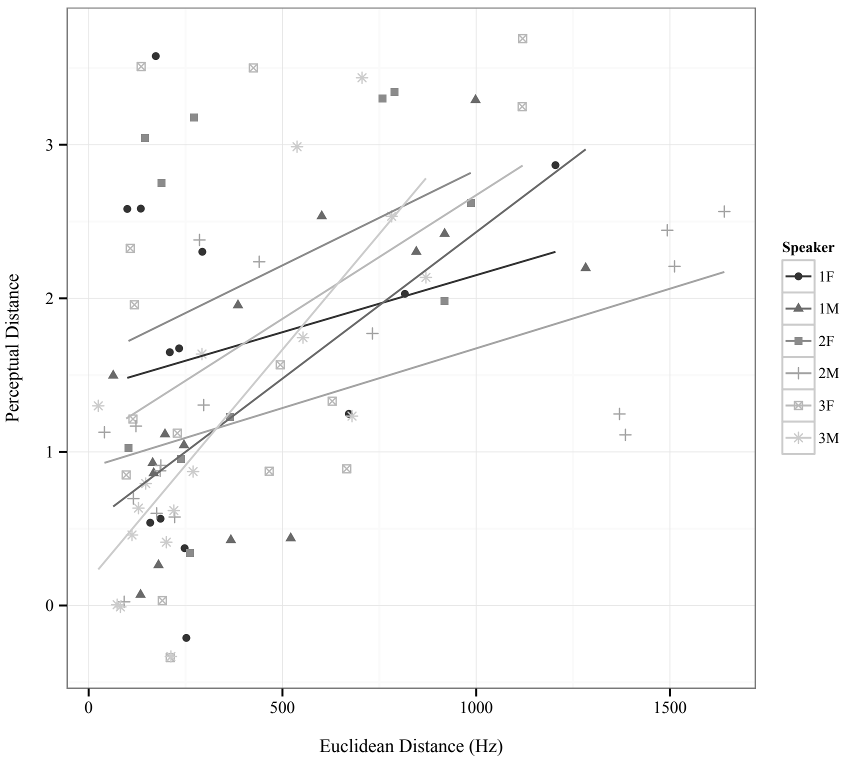
No significant correlation was found between perceptual similarity and acoustic similarity for words in the /pe/, /pi/, /po/ groups (one-tailed t test, $r = 0.69$, p

Table 4.11: Comparison of mean Euclidean distances in three dimensions (F_1 to F_3) with perceptual distances for male and female speakers, arranged by perceptual distance in ascending order for male speakers. Perceptual distances increase as Euclidean distance at vocalic onset increases ($r = 0.82$, $p < 0.05$ for males; $r = 0.49$, $p < 0.05$ for females). Vowel-medially the correlation is also significant ($r = 0.80$, $p < 0.01$. Data from Tables 4.11, 4.5, and 4.6.

CV_i—CV_j	Distance at voicing onset (Hz), male speakers	Perceptual distance, male speakers	Distance at voicing onset (Hz), female speakers	Perceptual distance, female speakers
po—ko	136	0.01	150	0.15
to—ko	95	0.53	366	2.19
po—kwo	252	0.60	202	0.72
ko—kwo	176	0.65	180	0.45
po—to	83	0.83	515	1.70
to—kwo	229	0.87	444	1.83
te—pe	195	0.90	155	2.45
pi—ti	120	0.96	288	3.49
te—ke	83	0.98	66	2.80
pi—ki	72	1.11	275	0.93
ti—ki	118	1.70	117	1.85
kwe—pe	967	1.81	1063	–
ke—kwe	1123	1.81	1178	3.42
pi—kwi	931	2.02	941	2.01
te—kwe	1108	2.48	1115	4.90
ti—kwi	989	2.64	937	–
ki—kwi	1001	2.88	823	2.35

< 0.1; see Table 4.11), suggesting that listeners may rely more on other cues to differentiate the /pe/, /pi/, /po/ words from those beginning with other obstruents. This may be corroborated by the relative typological likelihood of /p/ to develop into other labial obstruents such /f/ and /ɸ/, as in e.g. Germanic languages (McDorman, 1999). The lack of a significant correlation between acoustic and perceptual similarity for /pe/, /pi/, /po/ words is supported by the indication provided by Figure 4.5 that the relative acoustic similarity of individual speakers' /po/ and /kwo/ tokens did not relate to their likelihood of perceptual confusion, unlike for /ko/ and /kwo/ words.

Fig. 4.3: Plot of perceptual distance and three-dimensional Euclidean distance for all input/response combinations, divided by speaker.

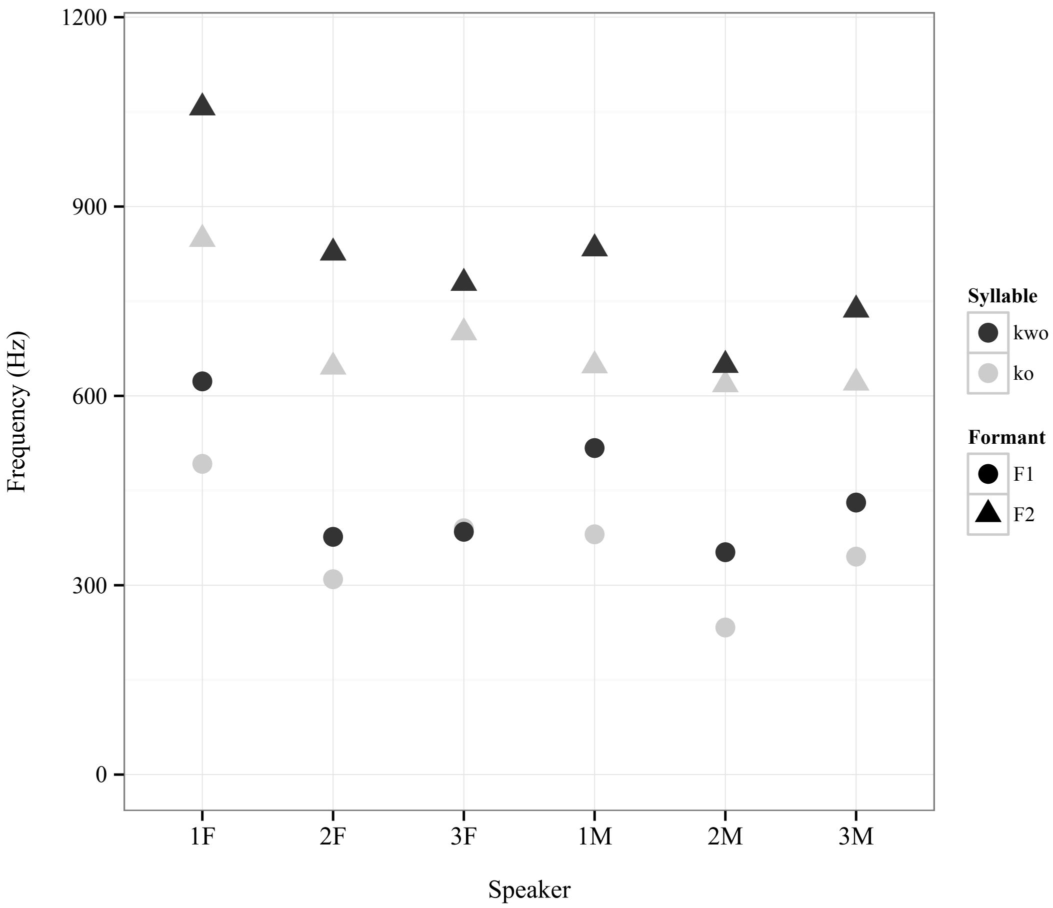


Variation between speakers

In section §3.5.3 close examination of Speaker 2F's tokens indicated that, at the onset of voicing, the difference between the first formant frequency of /kwo/ and that of /ko/ was greater than for any other speakers. The same is true of the second formant frequency, indicating that for Speaker 2F /kwo/ and /ko/ are more acoustically distinct than they are for other speakers (Figure 4.4). This could have rendered /kwo/ words more identifiable in relation to /ko/ words for this speaker, given that the former is characterized by lower frequency energy. There was a significant negative correlation (r = -0.92, $p < 0.05$) between acoustic distance of /kwo/ and /ko/ F_1 frequencies and rates of /kwo/ to /ko/ confusion (Table 4.5). The correlation between F_2 differences and confusion rates was also strong (r = -0.71) but not significant.

No effect of acoustic similarity was noted for confusion of /kwo/ and /po/ tokens (Figure 4.5). Speakers whose tokens elicited a zero rate of /kwo/ to /po/ confusion (speakers 1F, 2F, and 2M, see Tables 4.5 and 4.6) did not produce /kwo/ stimuli that were acoustically more similar to /po/ stimuli, unlike in the /kwo///ko/ example in Figure 4.4. Significant differences in formant transition frequencies exposed in 3.5.4 did not coincide meaningfully with increased likelihoods of perceptual confusion on a speaker-by-speaker basis and there was no correlation between acoustic similarity and perceptual confusion rates for each speaker's tokens. Two potential explanations come to mind: on the one hand, that the sample size was simply not large enough for meaningful differences between different stops to manifest ($n = 3$ per speaker) indicating that a definite conclusion is best postponed until an increased inventory of stimulus tokens is available; or on the other hand, the phenomenon noted in Section 4.4 (namely, the correlation between acoustic and perceptual similarity for words in the pV categories is neither strong nor significant) is behind the discrepancy, meaning that listeners are less reliant on acoustic cues to differentiate /po/ tokens both in general and in the context of comparison with /kwo/. This could conceivably

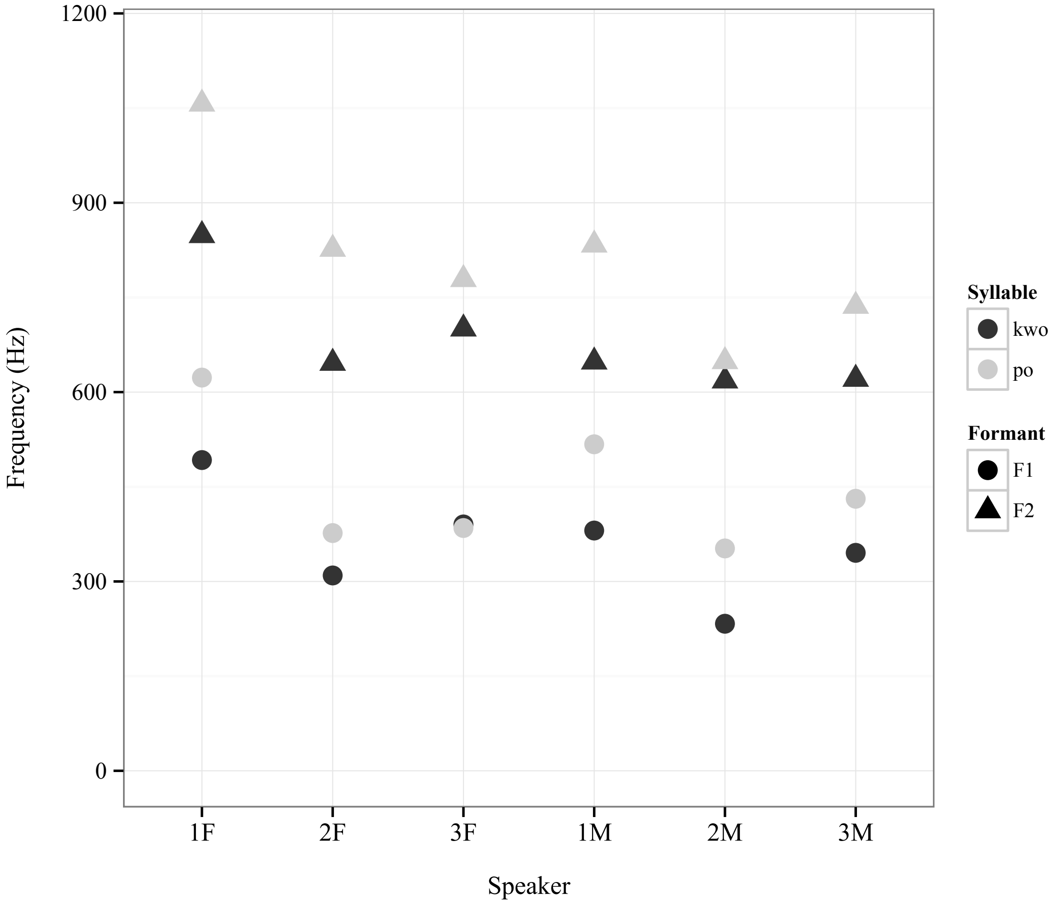
Fig. 4.4: Chart showing relative mean F_1 and F_2 values for /ko/ and /kwo/ tokens from each speaker, in order of decresasing acoustic distance between /ko/ and /kwo/. Note that the difference between the formant frequencies of /kwo/ and /ko/ words is comparatively greater for Speaker 2F, possibly accounting for the low levels of confusion between /kwo/ and /ko/ observed for this speaker's tokens (see Table 4.5).



lead speakers to be less motivated to articulate differences between /kwo/ and /po/ by the experiential knowledge that listeners rely on other cues to differentiate /po/ from other stops, although more articulatory data would be required to investigate this possibility further.

The comparatively high second formant frequency of /kwi/ and /kwe/ words produced by speakers 3F and 3M (noted in §3.5.3) did not appear to lead to higher

Fig. 4.5: Chart showing relative mean F_1 and F_2 values for /po/ and /kwo/ tokens from each speaker. Note that there is no apparent relationship between acoustic similarity of /kwo/ and /po/, which differs by speaker, and perceptual confusion rate of that speaker's tokens.



levels of confusion between /kwi/, /kwe/, and other word groups for those speakers (Tables 4.5 and 4.6). Likewise, Speaker 3F's exceptionally high F_2 frequencies for /ke/ words did not seem to affect perception in any meaningful way.

Additional observations

The most consistently strong perceptual bias (i.e. a bias which operated in the same direction in all vocalic environments) was that of confusion between words in the /ke/, /ki/, /ko/ and /te/, /ti/, /to/ groups; using pooled speaker data again, participants were almost twice as likely (1.97:1) to misidentify a /te/, /ti/, /to/ stimulus as a /ke/, /ki/, /ko/ word as vice versa. This pattern is strongly reminiscent of natural sound change patterns described by Blust (1990, 2004) and Blevins (2004); Blust identified twenty independent historical instances of $t < k$ change in forty-three languages (although most were of the Austronesian language family) and a further two were described by Blevins. Neither author identifies any corresponding examples of $k < t$ changes. In all these examples the change was preceded by the loss of *k* as a phonologically contrasting segment, leading to a period of free variation and/or a shift to *k* (free variation of *t* and *k* is found in Hawaiian, Hall 2010, p. 1).

The strongest perceptual bias was that between /pi/ and /ki/ words; /pi/ words were over four times as likely to be identified as /ki/ words than vice versa (0.40 compared to 0.09). I have not been able to find evidence of typologically common sound changes that correspond to this perceptual bias. Words in the /pe/, /pi/, /po/ groups were the most frequently misperceived of all stimuli, but *p* is amongst those sounds that change least in early Indo-European languages (data from Fortson (2010)). On the other hand, *p* is characterized by perhaps the most obvious of visual perceptual cues: complete lip closure, which was unavailable to participants in this experiment. McGuire & Babel (2012) conducted audio-visual perception experiments to investigate perceptual biases in [θ]/[f] confusion and found that consistency of visual cues was partly responsible for high rates of identification of [f] stimuli, presumably due to the presence of labial articulation. Similarly, Traunmüller & Öhrström (2007) found that experimental subjects perceived lip-rounding more readily by eye than by ear (although the experiment was conducted using changes in vowel quality rather than stop place articulation). Investigations in the McGurk effect (in which participants are exposed to non-matching visual and auditory stimuli) by Hardison (1999) found that perception of labial stops was most influenced by visual cues. Thus it seems plausible that the absence of visual cues in this perception experiment contributed to relatively high rates of /p/ confusion.

Lexical usage frequency effects

Usage frequency of lexical items has been shown to influence perception (e.g. Hay et al. 2004; Ferreira et al. 1996) suggesting that the use of natural words in this experiment may introduce bias into the subjects' selections. Exemplar phonology incorporates usage data not only from the analysis of phonetic patterns but from rate of recurrence of specific discourse models, lexical items etc. to provide explanations for linguistic phenomena with a limited recourse to hierarchical structure (Bybee, 2001). This approach is radically different from earlier rule-based theories, in which simplification of stored information was seen as necessary to produce maximally efficient rules or constraints (Chomsky, 1995; Gahl & Yu, 2006). Pierrehumbert (2001) suggests that stored exemplars form 'clouds' representing phonological categories and incorporating the physical and auditory correlates of the category, as well as experienced variation; the number and temporal closeness of specific variants is relevant in determining output.

A detailed analysis of the effects of lexical frequency on participants' misperception patterns was complicated by the lack of an accurate method of estimating lexical usage frequency for the sample population; existing linguistic corpora do not reflect the contemporary usage of neologism *quorn*, and usage of *quad* for the sample population (all of whom were connected to the University of Oxford) is likely to differ significantly from the norm. However, some analysis was possible using Google hits, which represent the frequency of occurrence of a given search term on the internet, i.e. the world's largest corpus of texts and one which includes many examples of spontaneous, contemporary written communication as well as literary and commercial texts. To investigate the possibility of participants being better attuned to higher-frequency words and therefore more likely to accurately identify them, total numbers of Google hits for each word were compared to the *d'* values calculated in this experiment (Table 4.9, in which *d'* values are collated by onset cluster).

No correlation was found between hits—scaled logarithmically—and perceptibility ($r = 0.28$, $p < 0.01$). To assess the relationship of number of Google hits to selection bias, hits were compared to the false alarm rate for each word in the experiment, i.e. the rate at which participants selected a word that was not the stimulus (biases in favour of specific words might reasonably be expected to manifest themselves in this data). There was no correlation between Google hits and false alarm rate.

Table 4.12: Table showing *d'* values and numbers of Google hits for the stimulus words used in Experiment 1. Some *d'* values are blank due to zero values in the equation used to calculate *d'*.

Word	*d'*	Hits (m)	$\log_{10}$ Hits
take	3.03	12550	10.099
test	2.952	6350	9.803
quick	3.89	4820	9.683
port		3290	9.517
tip	4.082	2750	9.44
till		2450	9.389
kill		2030	9.307
cake	3.285	1400	9.146
kick		1120	9.049
quest	3.949	773	8.888
taught		686	8.836
pit		641	8.807
quit		638	8.805
porn	1.902	546	8.737
kit	3.641	464	8.667
quad	2.773	462	8.665
pick	2.63	380	8.58
tit	3.68	280	8.447
torn	2.53	267	8.427
pod	1.389	215	8.332
pest	2.999	192	8.283
tick		188	8.274
pip		185	8.267
tod	2.797	166	8.220
quake	3.57	157	8.196
cod	1.709	140	8.147
quart		98	7.991
corn	1.519	68	7.831
pill	2.864	42	7.618
quill	4.220	32	7.498
kip	3.908	19	7.199
quip		16	7.199
quorn	1.793	0.8	5.893

4.5 Conclusion

This experiment did not reveal a stong likelihood for listeners to misperceive /kw/ as /p/, although there was a tendency for listeners to identify /po/ as /kwo/. This calls into question the legitimacy of the perceptual confusion approach, in which perceptual biases may lead directly to sound changes; I have not found any instances of historical sound changes in which labials became labialized velars. It is possible that a tendency to confuse /kw/ and /p/ does not lead to sound change for reasons discussed by Chang et al. (2001), which uses concepts from visual perception psychology to suggest that subjects may remove elements during reanalysis but not add them; in this instance, despite a perceptual bias towards confusion, subjects may remove a velar articulation but would not add it. In addition to these concerns, the difficulty of actually eliciting misperception (most tokens were identified correctly even with a severely distorted stimulus signal) begs the question of how realistic misperception is as a precursor to historical phonological change. However, auditory confusion experiments do not incorporate the full range of information available to listeners; the high confusion rate of /p/ in this experiment and the evidence described in §4.4, suggesting that visual cues are comparatively salient in the identification of labial sounds, indicates that perceptual biases concerning labials would differ if visual cues were available to participants.

Interestingly, /ko/ and /po/ were the most perceptually similar sequences included in the experiment, and were more similar to one another than /kwo/ and /po/ for both male and female speakers. Given the high level of acoustic and perceptual similarity between /kwo/ and /ko/, it is not clear why /ko/ is highly confusable with /po/ but /kwo/ is not. It may be that the longer VOT of /kwo/ tokens (Fig. 3.12) is used as a cue to distinguish /kwo/ from /po/, although this explanation is not entirely satisfactory because /ko/ has a VOT closer to /kwo/ than to /po/.

The acoustic similarity between /kwo/ and /ko/ described in Chapter 3 led to perceptual confusion in this experiment; there was a strong tendency for participants to confuse /kwo/ with /ko/. As predicted in Chapter 3, sequences of /kw/ followed by front vowels were most resistant to perceptual confusion for all speakers and listeners, probably due to the distinctive steeply rising F_2 transition. The most easily identifiable tokens were those featuring /t/ before front vowels, possibly due to the dominant high frequency peaks that appeared in the burst spectra of those tokens.

5

Phonetic explanations for the Greek sound changes

5.1 Introduction

In the previous chapters, conditioned sound changes involving labialized velars in Indo-European and Ancient Greek were introduced, and acoustic and perceptual evidence from modern languages was presented to investigate possible phonetic causes for these sound changes. In this chapter, evidence from these sources is brought together and combined with typological evidence from outside Greek and Indo-European to consider a phonetically motivated account for each conditioned outcome in Greek.

Each conditioned outcome will be considered separately. First, Greek developments of labialized velars to labials are discussed; second, Greek developments of labialized velars to coronals are discussed; third, Greek developments of contrastively labialized velars to non-contrastively labialized velars are discussed. For each of the three sound changes, typological evidence from other language families is introduced and analyzed to determine the relevance of phonetic accounts, conclusions from the acoustic analysis in Chapter 3 are discussed with reference to the Greek sound changes, and lastly the role of perception is considered using the perceptual data from Chapter 4.

5.2 Developments to labials

5.2.1 Introduction

In Chapter 1, arguments were presented that the development of labialized velars to labials may be a phonetically abrupt change prompted by perceptual similarity between the two sounds (Ohala, 1993; Garrett & Johnson, 2013). However, in Chapter 4 a perceptual confusion experiment using British English tokens revealed that there was no significant tendency for listeners to mistake labialized velars for labials in any of the vocalic environments tested, despite severe masking of the speech signal by noise intended to provoke misperception. This section presents evidence from multiple language families in which labialized velars developed to labials, providing support for a physiological or phonetic motivation; the plausibility of the perceptual confusion account is then assessed with reference to the findings from acoustic and perceptual experiments in earlier chapters, before alternative explanations are explored.

5.2.2 Typological parallels: labialized velars to labials

There are numerous examples of labialized velars developing into labials in geographically, temporally, and linguistically diverse settings. This section presents examples of these sound changes, beginning with those attested within the Indo-European language family and then describing similar changes in other language families.

Ancient Greek A discussion of Greek outcomes was presented in Chapter 2.

Scots The following sound changes involve velar fricatives with secondary labialization developing into labial fricatives (Examples 5.1–5.3, from Garrett & Johnson 2013).

(5.1) OEng. /xwaː/ to Buchan Scots /faː/ 'who'.

(5.2) OEng. /xwiːt/ to Buchan Scots /fəjt/ 'white'.

(5.3) OEng. /xwonne/ to Buchan Scots /fan/ 'when'.

Uto-Aztecan Some Uto-Aztecan languages show a labial reflex /b/ in positions where other languages have inherited /kw/, as in Example 5.4.

(5.4) O'odham *bahi*, Mountain Pima *bahi* 'tail', cf. Northern Paiute *kwasi* 'tail', Hopi *kwasi* 'penis'.

Acula, Tuxtlas, and Pajapan varieties of Nahuatl show labial outcomes [b] and [ƀ], in variation with one another; Jalupa Nahuatl has [bʷ] in these positions (Monzón & Seneff, 1984; García de León, 1967). The development of /b/ from /kw/ in Tuxtlas is described as being conditioned by a following back vowel /ɑ/ by Hasler (1958). (Monzón & Seneff, 1984, p. 456n) report that many local dialects of Nahuatl have a sound which they transcribe [w̸] and describe as "characterized by the parallel position of the lips (as for [ƀ]) and a velar onglide–no friction has been noted" (the symbol [ƀ] represents a voiced bilabial fricative [β]). This sound might be transcribed in IPA as [ɰʷ] or [ɣ͡β] (Pullum & Ladusaw 1996, p. 163).

Examples of Nahuatl reflexes with /b/ and /p/ are provided in Examples 5.5–5.12 (Stubbs, 1995). Shaul (2010, p. 271) writes that the Eudeve /b/ reflex from $*k^w$ has developed to voiceless /p/.

(5.5) Eu. *bowá* 'invite', from proposed $*k^w$*awe* 'invite', cf. Cupeño k^w*awe* 'call, invite'.

(5.6) Eu. *puse-* 'make dust', from proposed $*k^w$*usi* 'dust', cf. Np. *kusi-bɨ* 'dust'.

(5.7) Eu. *sébor* ‘fly’, from proposed *sek^wori* ‘fly’, cf. Cáhita *kuŋ-sex^{w}et* ‘bumblebee’, Mayo *sé’ebori* ‘fly’, Yaqui *sé’ebo’i* ‘fly’, Classical Nahuatl *šiiko’-*.

(5.8) Eu. *purúce* ‘hang’, from proposed *k^wut* ‘hang’, cf. Hopi *kolca* ‘shelf’.

(5.9) Eu. *boc* ‘grandfather’, from proposed *k^woci* ‘paternal grandfather’, cf. Wr. *wocí* ‘grandfather’, Yq. *haboi* ‘paternal grandfather’; the reconstruction could instead be **poci*, but Classical Nahuatl *kool-li* supports the labialized velar (Stubbs, 2011, p. 200).

(5.10) Eu. *zápa-n* ‘grab, seize’, from proposed *cak^{w}V* ‘grasp’, cf. Cupeño *cák^{w}e* ‘catch’.

(5.11) Eu. *bá* ‘food’, cf. Jova *ca* ‘food’ and Op. *guaka* ‘food’, from the Lord’s Prayer transcriptions in Shaul (2010, p. 266–267).

(5.12) Proto-Nahuatl *k^wi* ‘to catch, hold’, cf. Mayo *a’a b^{w}i-sse* ‘that which is caught’ (Dakin, 1982).

These examples are particularly interesting because Proto-Nahuatl is not generally reconstructed with a voiced labial /b/ (Stubbs, 2011; Steele, 1979), making the development of /b/ from *k^w structurally inconsistent; the sound change introduces a lone voiced stop that has no other contrasting voiced stops. Spanish is spoken widely in the areas studied and does feature voiced stops, including /b/. However, Monzón & Seneff (1984) argue against the influence of Spanish in determining the course of this sound change. Despite the presence of some completely bilingual speakers of dialects whose outcome stops have already acquired bilabiality, the influence of Spanish [b] and [β] (which occur in conditioned allophonic distribution) is not apparent in the earlier stages of the change; conditioned allophonic variation is not noted in most other speakers (Monzón & Seneff, 1984). The formation of voicing might be thought plausible in intervocalic contexts (e.g. Examples 5.7 and 5.9) but word-initially it seems less likely to be a spontaneous, phonetically-motivated development.

Muskogean Labialized velars are reconstructed for Proto-Muskogean based on correspondence sets such as those in Examples 5.13-5.15 (Booker, 1993, p. 405–410).

(5.13) Proto-Muskogean **kwixi* ‘fig’, giving Choctaw *bíhi* ‘fig’, Creek/Seminole *ki* ‘fig’.

(5.14) Proto-Muskogean **sakkwa* ‘arm’, giving Choctaw *šakba* ‘arm’, Creek/Seminole *sakpa*.

(5.15) Proto-Muskogean **yokwala* ‘behind’, giving Choctaw *obala* ‘behind’, Creek/Seminole *yópa* ‘behind’.

Zapotec Some Zapotec varieties show labials in positions where others show inherited labialized velars, as in Example 5.16 (Ohala, 1989).

(5.16) Matatlan Zapotec *kwan*, cf. Isthmus Zapotec *pa*.

Caddoan Examples 5.17–5.18 (from Taylor 1963) demonstrate historic alternation between labial stops, labialized velar stops, and labial-velar approximants in Caddoan languages.

(5.17) Proto-Northern-Caddoan **kʷiˑt* ‘man’, Pawnee *piˑta*, Wichita *awáˑhiskʷiˑ*cʔ ‘Pawnee man’ (cf. *awáˑhiˑ* ‘Pawnee’).

(5.18) Proto-Northern-Caddoan *-*waˑwa* ‘eat’, Pawnee *tut-pá-wa* ‘I eat’, Wichita *tackʷáˑ*w*aˀàs* ‘I’m eating’.

Bantu Proto-Bantu velars with coarticulatory labialization resulting from proximity to a following /u/ show labial outcomes in West Teke, as in Example 5.19 (Ohala, 1989).

(5.19) Proto-Bantu *-*kumu* giving West Teke *pfumu* ‘chief’.

In Sawabantu, an inherited prefix **ku*- developed into *kw* before vowels, and subsequently to *kp* in Western Sawabantu (Cahill, 1999):

(5.20) Eastern Sawabantu *kwalé*, Western Sawabantu *kpaé* ‘partridge’.

(5.21) Eastern Sawabantu *kwatá*, Western Sawabantu *kpatá* ‘sword’.

(5.22) Eastern Sawabantu *kwédí*, Western Sawabantu *kpélí* ‘death’.

Hombert et al. (1989) seek to establish a genetic relationship between Fang and other Bantu languages, and propose a direction of change from inherited *k^w and *g^w to /k͡p/ and /ɡ͡b/ on the basis of correspondences such as the following:

(5.23) Proto-Bantu **kùàdí* to Fang /òkpàá/ ‘partridge’.

(5.24) Proto-Bantu **búd* to Fang /áɡbéle/ ‘to break’.

(5.25) Proto-Bantu **bùd* to Fang /áɡbɛ̀ɲ/ ‘to rot’.

Interestingly, Hombert et al. (1989) also provide examples in which similar historic sequences become labio-palatalized velars when the following Fang vowel is a front vowel:

(5.26) Proto-Bantu **kóddá* gives Fang /kɥɛ́ɲ/ ‘snail’.

(5.27) Proto-Bantu **kúúd* gives Fang /ákɥɪɲ/ ‘exit’.

(5.28) Proto-Bantu **gùdú* gives Fang /ŋɡɥɪ̂ɲ/ ‘pig’ (cf. e.g. **gòmà* to /ŋɡɔ̀m/ ‘pig’).

Hombert et al. do not discuss this potential conditioned development, but it is of possible relevance to the similar Greek developments.

Aghem Aghem is a Grassfields Bantu language in which /kp/ and /gb/ developed from labial stops with secondary labialization, **pw* and **bw* (Cahill, 1999).

Luganda Herbert (1975) observes that the Luganda language of Uganda has surface forms [f] and [v] resulting from earlier sequences of /ku/ and /gu/. It is argued that this is due to the lowering of F$_2$ transitions of velar consonants under the influence of an adjacent round vowel, rendering them more acoustically similar to labials (p. 116).

Niger-Congo Noni belongs to the Niger-Congo family of languages and is spoken in Cameroon. Velar stops with double and secondary labialization alternate phonetically; see Examples 5.29-5.30 (Cahill, 1999).

(5.29) *kwen* or *kpen* 'firewood'.

(5.30) *gwéé* or *gbéé* 'hundred'.

Elsewhere there is cross-linguistic variation between /k͡p/ and /kw/ in Niger-Congo languages; see examples 5.31-5.32 (Blench, 2004).

(5.31) **kpa* 'to sweep', gives Yoruba *gba*, Ahan *kpàá*, Urhobo *è-xwérè*, Onica *kʷa*.

(5.32) **a-kpa-* 'bridge', gives Yoruba *afárá*, Edo *agbadi*, Igbo *àkwà*, Idoma *akpa*, Eggon *apa*, Takum *taba*.

Within the Nupoid subgroup, Nupe and Gbari show labial-velars and labialized velars respectively (Westermann & Ward, 1933, p. 108).

(5.33) Nupe *kpe* and Gbari *kpe* 'cover'.

(5.34) Nupe *kwa* and Gbari *kpa* 'feed'.

(5.35) Nupe *gwa* and Gbari *gpa* 'humble'.

(5.36) Nupe *gwere* and Gbari *gbere* 'red'.

Gbe The Gbe languages of eastern Ghana and western Nigeria show divergent developments of inherited *xʷ to a voiceless bilabial fricative /ɸ/ in the Ewe subgroup and /p/ in the Gen subgroup. Examples 5.37-5.42 are from Capo (1991, p. 109–110); note delabialization adjacent to a back round vowel in some languages, see example 5.39.

(5.37) Proto-Gbe **aχʷá* 'outcry', gives Ewe *aɸá*, Gen *apá*, Fon *aχʷá*, Ajá *aχʷá*.

(5.38) Proto-Gbe *-*χʷã* 'sniffling', gives Ewe -*ɸã*, Gen *epã*, Fon -*χʷã*, Ajá -*χʷã*.

(5.39) Proto-Gbe *χʷ*ú* ‘to dry’, gives Ewe ɸ*ú*, Gen *pú*, Fon χ*ú*, Ajá χ*ú*

For the labialized uvular fricative *ʁʷ, the outcome in Ewe is a labiodental approximant /ʋ/ (Examples 5.40-5.39; once again there is delabialization adjacent to a back round vowel, see Example 5.39).

(5.40) Proto-Gbe *ʁʷe ‘be small’, gives Eve ʋə, ʋɛ, ʋe, Gen ʁʷe, Fon ʁʷe.

(5.41) Proto-Gbe *ʁʷ*ã* ‘to move’, gives Ewe ʋ*ã*, Gen ʁʷ*ã*, Fon ʁʷ*ã*.

(5.42) Proto-Gbe *-ʁʷ*ũ* ‘blood’, gives Vhe -ʋ*u*, Gen eʁ*ũ*, Fon -ɾ*ũ*.

The labialized palatal approximant /ɥ/ alternates with /w/ before front vowels in all Gbe dialects (Capo, 1991, p. 56) e.g. /kɥi, kui, kwi/ ‘take it’.

Bari Bari is a Nilotic language spoken in Sudan. Westermann & Ward (1933, p. 109) describes the Kakwa dialect of Bari, in which Bari /kw/ appears as /kp/.

(5.43) Bari *kwen* and Kakwa *kpen* ‘birds’.

(5.44) Bari *gwagw*ɛ and Kakwa *gbagb*ɛ ‘wild cat’.

Caucasian languages Catford (1977c, p. 292) and Smeets (1986, p. 286) describe sound changes in the Circassian group of languages in which velar fricatives with secondary labialization become labiodental fricatives:

(5.45) Proto-Circassian *x^w to Adyghe /f/, cf. Kabardian /xʷ/.

Finnish The velar fricative /ɣ/ is realised as [v] in positions between high round vowels /u/ and /y/ (Jacques, 1990, p. 49).

Hausa Zima (1974, p. 129) observes that Hausa /f/ is realised by speakers in the town of Dogondutchi as [xʷ], see example 5.46.

(5.46) Hausa *Fáatimá* and Dogondutchi *Chʷátʃimá*, personal name.

Songhay Zima (1985) describes allophonic variation between [k͡p] and [kw], with [k͡p] available as an 'optional variant' of [kw] to speakers of the Djougou Dendi dialect of Songhay (p. 98) .

5.2.3 Perceptual confusion account

Neither the acoustic analysis in Chapter 3 nor the perceptual confusion experiment in Chapter 4 provide strong support for the perceptual confusion account as proposed by Ohala (1990b, 1993) and discussed in Garrett & Johnson (2013). It was demonstrated that there was no particular likelihood for listeners of British English to perceive /kw/ as /p/ in any vocalic environment. However, as discussed in §3.7, there is evidence to suggest that the timing of the velar and labial articulations in labialized velars differs cross-linguistically, although it was not tested if this affects perception. For Western Zapotec, the acoustic effects of labialization are more prominent at the burst than in the formant frequencies of the following vocoid; for British English, labialization affects the vocoid more prominently, exposing acoustic evidence of coarticulation of the velar articulation. It is possible that for an articulation more like the Zapotec, in which the acoustic effects of lip rounding are detectable simultaneously with the velar burst, and are less prominent at the onset of voicing, a perceptual confusion analysis with labials is more likely. On the other hand, the burst spectra for Zapotec and British English /kw/ did not display any particular acoustic similarity to those of British English /p/ in any environment; for /kw/ bursts there was a dominant low frequency peak (for Zapotec, Figs. 3.15, 3.17 and 3.19; for British English, Figs. 3.5b, 3.6d, 3.7d and 3.8d), but for /p/ bursts the spectrum was diffuse (Figs. 3.6a, 3.5a, and 3.8a). Also, similar sound changes have been described for dialects of English (see the Buchan Scots examples in §5.2.1, in which /xw/ becomes /f/), suggesting that labialized velars with sequential lip-rounding are susceptible to developing into labials. Although the case of labial-velars is not directly comparable to labialized velars, Connell (1994) suggests that the reason for the typological likelihood for labial-velars to develop into labials, rather than velars, lies in the asynchrony of the gestures and the possibility that the labial release occurring later makes it more perceptually salient. Thus a lesser degree of articulatory overlap may help rather than hinder a sound change of /kw/ to /p/. Interestingly, Ponelis (1974, p. 39–41) argues that segmentation—or reanalysis of labialization as a "distinct element" following the first constriction—is a necessary prerequisite of developments of e.g. /kw/ to /kp/. Goldstein (1995) develops a similar idea, describing how articulatory overlap creates aerodynamic conditions which prompt category reassignment of specific gestures for changes of English /oʊx/ to /ɒf/ (as in e.g. *cough*).

The difficulty of eliciting misperceptions amongst the participants in the perceptual confusion experiment described here calls into question the theoretical validity of perceptual confusion accounts. Even with severe distortion of the speech signal participants were able to correctly identify the majority of tokens. Thus a misparsing account, whilst it may be appropriate in cases where labialized velars lose contrastive labialization adjacent to rounded vowels (§5.4) due to a combination of acoustic and articulatory similarity, may be less plausible for this sound change. On the other hand, participants were (presumably) devoting their full attention to the discrimination task, in contrast with normal speech interaction where distractions are generally present.

5.2.4 Intermediate stages

An alternative explanation involves articulatory changes via stops with double articulation, possibly motivated by perceptual biases or misparsing. Several examples presented in §5.2.2 show sound changes of labialized velars to doubly articulated labial-velars; multiple examples from languages of the Niger-Congo family show cross-dialectal variation between /k͡p/ and /kw/, and /g͡b/ and /gw/. Dialect variation amongst speakers of Nahuatl shows development of /kw/ to double articulated approximants [ɣ͡β] and [w]. In turn, some of these labial-velar stops became plain labials, as in the Uto-Aztecan alternations observed by Monzón & Seneff (1984) and in the Niger-Congo cognates presented in Example 5.32. These examples suggest that there is a typological likelihood for labialized velars to alternate with labial-velars and plain labials, suggesting it is plausible that similar sound changes could have occurred in Ancient Greek. However, the extent to which the two types of sound—labialized velars and labial-velars—differ in articulation and acoustics is not clear. The distinction between secondary and double articulation may simply be a matter of varying degrees of constriction, as hypothesised by Catford, but factors such as length of lip protrusion and degree of rounding may affect the phonetic or phonological realisation of the sounds.

Ponelis (1974, p. 39) describes a multiple stage progression of articulatory change from /gw/ to /b/ through 'glide narrowing' in a process which yields an intermediate stage [gβ]. The next stage of this process is implied by evidence such as that offered by Connell (1994), who describes instances of Ibibio labial-velar stops in which the velar closure was weak or incomplete relative to the plain velar closure (p. 472). Thus there is evidence for intermediate stages of development involving gradual closure of the labial constriction, followed by weakening and eventual redundancy of the velar closure.

The reason for alternations between labialized velars and labial-velars may also be due to their similar acoustic characteristics, as well as their articulatory similarity. Formant transitions of labial-velar stops were described in §3.2.3 as having a lower F_2 frequency and a steeper transition into a following vowel than plain labials; these were also the acoustic characteristics which characterized labialized velar stops in the languages analysed in Chapter 3, indicating an acoustic similarity between labial-velars and labialized velars.

5.2.5 Conclusions concerning developments to labials

The confusion experiment in Chapter 4 provided little evidence to suggest that perceptual confusion could provide a motivation for changes of labialized velars to labials. The observation of gesture timing differences between British English and Western Zapotec labialized velar sounds indicates that the degree to which labialization overlaps with the velar burst is language-specific, or differs with the phonological status of the sounds. Thus, for languages in which labialization and the velar constriction overlap more fully than in English, there may be a greater likelihood of perceptual confusion with labial stops. On the other hand, when compared to British English labials the Western Zapotec labialized velars are not spectrally similar, having a dominant low frequency peak whilst the labials have peaks generally in line with the formant frequencies of the following vowels due to the relative coarticulatory freedom of the tongue. In conclusion, there are few reasons to suppose that an acoustic misperception account involving an articulatorily immediate transition from /kw/ to /p/ is the most plausible for this sound change. Alternative explanations have focussed on an articulatory progression involving a gradual increase in the degree of constriction at the lips, followed by loss of the velar closure; this analysis is supported by typological evidence from several language families. However, it is possible that these two approaches might be combined. A potentially high level of acoustic similarity between /kw/ and /k͡p/, coupled with evidence from numerous languages indicating intermediate forms along an articulatory continuum from labialized velars to labial-velars to labials, suggests an account involving both articulatory and perceptual factors. Although there is not a great deal of literature on the subject,Ladefoged (1964) and Connell (1994) describe the acoustic effects of labial-velars on adjacent vowels, which are similar to those of the labialized velars analysed here in Chapter 3. Thus it is feasible that Ancient Greek may have selected variants with increased labial constriction without ever passing through a phase in which /k͡p/ contrasted with /kw/; the two may have appeared in complementary distribution depending on vowel environment. It seems there is a typological basis for this type of alternation to take place, and there is an indication that there could be an acoustic basis also, but the matter requires further investigation.

5.3 Developments to coronals

5.3.1 Introduction

In §2.5.5, inscriptional evidence and glosses from Arcadian and Lycian were presented to suggest an intermediate stage between the Indo-European labialized velar and the coronal outcomes in Greek before front vowels. In Chapter 3, burst spectra and formant frequencies of labialized velar and coronal stops in front vowel environments were compared; there was little acoustic similarity, with labialized velars marked by a rapid dynamic increase in F_2 frequency during the formant transition and coronals distinguished by a prominent high frequency peak in the burst spectrum. In the following discussion, I present proposed intermediate stages from several philological sources and discuss their relative likelihood, taking into account the results from the acoustic and perceptual experiments described in Chapters 3 and 4.

Intermediate stages Evidence from two ancient languages—Lycian and Greek—suggests an intermediate stage of development for labialized velars in front vowel environments, at a stage at which their development had diverged from that of labialized velars in other phonetic environments but not yet merged with the coronal series. Evidence from Lycian inscriptions was presented in §2.2.3, and evidence from inscriptions and glosses in the Arcadian dialect of Greek was presented in §2.5.5. The implications of that evidence, and its interpretation by modern scholars, is discussed in this section.

In Lycian, as discussed in §2.2.3, the reflex of a labialized velar before a vowel /e/ is sometimes represented with a special character <τ> (see Examples 2.11 and 2.12). Philologists have suggested that this letter may represent a palatal /č/ or /ć/ (Pedersen, 1945, p. 13) or a coronal with secondary labialization, [t^w] (Melchert, 1994, p. 39–41). Pedersen's interpretation is something of an educated guess, and the interpretation as [t^w] may be based on the possible use of <τ> to transcribe the outcome of PIE *tw in Example 2.11; but the interpretation of this word as 'four' is doubtful, and Neumann (2007) notes that it may actually refer to a town. Melchert (2004) later retracted his earlier identification of the sound as /t^w/ and identified it as a palatal plosive /c/, but this suggestion is made on the basis that it 'seems a plausible transition sound' (p. 594). Thus there is little reliable evidence on which to draw conclusions about the phonetic attributes related to the sign <τ> in Lycian.

Rather more has been written about the developments in Ancient Greek. The consensus amongst specialists in Ancient Greek is that the development of labialized velars to coronals was articulatorily incremental (Lejeune, 1972; Risch, 1979; Brixhe, 1996; Bubeník, 1983). First, there is a concern that loss of one articulatory gesture (i.e. /kw/ changing to /k/) would have led to /k/ falling together with phonemes already present in Greek at the time. That is to say, loss of labialization to give a plain /k/, identical to the pre-existing Greek /k/, which then itself developed to /t/ cannot be satisfactory unless the other instances of /k/ in the language also showed that development; this is not the case. Second, inscriptional evidence in Arcadian shows a grapheme in the place of the reflex of a labialized velar before front vowels (§2.5.5); the sound represented was unlikely to have been a labialized velar (see discussion in §2.5.5) but was apparently different from the plain coronal outcome observed in Greek of the same geographical region some centuries later, and from the plain coronal inherited from Indo-European and represented with the Greek letter T in the same inscription (Duhoux, 2007). This has led many to assume an intermediate stage between *k^w and /t/. Beyond these starting assumptions, the motivation for these sound changes lacks general agreement, although philologists have suggested gradual processes of change by which they could have taken place:

$*k^{w} > *k^{y} > *k^{s} > t$ or $*s$

Lejeune (1972)

$*k^{w} > *k^{\grave{w}} > *kj > *k' > *t' > *t'\acute{s} > \acute{s} > s$

$*k^{w} > *k^{\grave{w}} > *kj > *k' > *t' > t$

Risch (1979)

$*k^{w} > *kj > *k' > *\underset{\frown}{t} > t$ or s

Brixhe (1996)

$*k^{w} > *k^{\text{'ɥ}}i > *\check{c}^{\text{ɥ}}i > *\check{c}i > t^{s}i > si$

Bubeník (1983)

The notation used differs between authors: [’] is used in some historical IE studies to indicate fronting, whilst [̑] indicates palatalization (Pullum & Ladusaw, 1996, p. 216, 230). Lejeune, Risch, and Brixhe suggest palatalization of the secondary articulation to [j] before any change affects the velar articulation. Bubeník suggests retention of labialization of the glide but addition of palatalization, giving a labio-palatal secondary articulation, with simultaneous fronting of the velar articulation. All four approaches first propose modification of the secondary articulation before or at the same time as fronting of the primary articulation, and all show an articulatory progression involving both palatalization and affrication at some stage. The suggestions from Lejeune, Risch, and Brixhe propose modification of [ʷ] first, and the first stage of development for all three may be summarized in IPA orthography thus:

$/\mathrm{k}^{\mathrm{w}}/ > /\mathrm{k}^{\text{ɥ}}/$

Bubeník’s account differs in that changes affect both the velar and labial-velar articulations simultaneously:

$/\mathrm{k}^{\mathrm{w}}/ > /\underset{+}{\mathrm{k}}^{\text{ɥ}}/$

Although these accounts focus on palatalization as a mechanism for the development, it is not clear how this could have taken place. Fronting of ‘plain’ velars as a coarticulatory mechanism in positions preceding front vowels, leading to historical sound change, is a well-attested linguistic phenomenon (Guion, 1996, 1998); however, whether or not front vowel environments have similar acoustic effects on velars that have secondary labialization has not been demonstrated.

Structural accounts Allen (1957) offers a comprehensive structural explanation of the sound changes affecting Greek labialized velars and relates the developments to phenomena occuring apparently concurrently in other labialized stops. Allen argues that the labial segment, rather than the velar segment, underwent palatalization first (hence [k^w] > [$k^ɥ$]) which caused the subsequent more palatal articulation of the velar. Utilizing the ‘empty space’ concept after Martinet, Allen suggests that the development of /s/ to /h/ left a structural gap in the phonological system which was filled by changes of /t(j)j/ and initial /tw/ into /s/ and, more importantly, /ti/ into /si/ (this despite the analogical retention of /s/ in verbal forms). This left an empty space for /ti/ which was duly filled by /kwi/.

As Allen admits, this theory suffers from the observation that the ‘gap’ left by the development of most instances of /s/ to /h/ was never completely empty; the usage frequency of /s/ was undoubtedly reduced, but whether or not such frequency reductions are typologically a genuine catalyst for other sounds to begin developing into the now less frequent sound is not clear. Also, for /kwi/ to “move into” the empty space previously occupied by /ti/, some mechanism needs to have existed by which /kwi/ is judged to be an acceptable combination of sounds to replace /ti/. Thus the discussion returns to the question of phonetic advancement from /kwi/ to /ti/.

5.3.2 Typological parallels: labials to coronals

Labialized velars to coronals

Historical sound changes of labialized velars to coronals are not typologically common, but examples do exist of synchronic and diachronic alternations between these two groups of sounds.

Chinese Reconstructed Ancient Chinese had an allophonic distinction between palatal affricates with secondary labialization, which occurred in front vowel environments, and velars with secondary labialization, which occurred in other positions (Karlgren, 1923).

(5.47) Modern Mandarin *qū*, IPA [tɕʰýˑ], 'bent' from Ancient Chinese kʰi̯wok.

(5.48) Modern Mandarin *juǎn*, IPA [tɕyɛ̰̀n], 'roll up' from Ancient Chinese ki̯ʷän.

(5.49) Modern Mandarin *guàng*, IPA [kwân], 'restaurant' from Ancient Chinese kuân.

Akan Labio-palatalization (simultaneous labial rounding and palatal constriction) occurs in complementary distribution with labialization in Akan (Westermann & Ward 1933, p, 106–107; Hall-Lew 2006). Synchronic phonological processes result in alternations as in Example 5.50.

(5.50) Asante Twi dialect, [kᶣia] 'farming' contrasts with [akʷʊa] 'servant'.

Labio-palatalization is not confined to velars; labio-palatal glides alternating with labial glides are observed after nasals, coronals, and labiodentals, see Examples 5.51-5.53 (other environments are provided in Hall-Lew 2006, p. 36).

(5.51) Akuapem Twi [mʷʊ̃ã] 'to dent', cf. [mᶣia] 'shut (mouth)'.

(5.52) Fante Akan [twei] 'to puncture', cf. [tᶣɪa] 'to join'.

(5.53) Asante Twi [ɛf^{ɥ}ia] *proper name*, cf. [f^{w}ʊa] ‘to add’.

Hall-Lew’s analysis focusses on the role of the phonological feature [Advanced Tongue Root] in vowel and stop harmony patterns in Akan, and outlines optimality constraints to determine the place of articulation of a secondary glide in specific vowel environments. Precise environments in which labio-palatalized glides appear in the surface forms differ by dialect, but it is worth mentioning that the labio-palatal glide does not occur before back rounded vowels in any of the examples cited in the paper and seems to be due to a convergence of anticipatory labialization and palatalization.

Niger-Congo Blench (2004) and Gerhardt (1983, p. 100) describe correspondence sets for Benue-Congo (a subgroup of the Niger-Congo family) languages in which an original labialized velar gives labial, labialized velar, palatal, and coronal outcomes (although Blench reconstructs an original labial-velar).

(5.54) East Benue-Congo **ikpi* or **ikwi* ‘rat’, gives Nupe *etsú*, Lopa *kyau*, Hasha *ikwi*, Alumu *ì-kwì*, Berom *cú*, and Hyam of Kwoi *kpyi*.

Nzema and Dagbani both have [t͡p, d͡b] as allophones of /k͡p, g͡b/ in positions preceding front vowels and [y] (Welmers, 1973, p. 46–47), suggesting coarticulatory fronting of the velar articulation despite the intervening labial release (Cahill, 1999, p. 158). Dagbani examples are provided below; the fronted articulations are more palatal than velar (Wilson & Bendor-Samuel, 1969, p. 58).

(5.55) Dagbani [tpi] ‘die’.

(5.56) Dagbani [dbi] ‘dig’.

Athabaskan Dialects of the Athabaskan Na-Dene language family show affricates developing into labialized velars, suggesting bidirectionality in sound changes of this kind. Cognates from dialects of Slavey are presented in Table 5.1. According to the standardized orthography in Rice (1989), <ddh> represents an interdental affricate [dð], <tth> represents an aspirated interdental affricate [tθ^{h}], and <tth’>

is an interdental ejective [tθ']. Flynn & Fulop (2008) present similar data. The proto-Athabaskan phonemic inventory has been revisited and revised since the publication of Rice's correspondence sets (Leer, 2005; Hargus, 2010) but the derivation of labialized velars from some type of dental or alveolar affricate remains accepted.

Table 5.1: Table of correspondences showing dental affricates corresponding to labialized velars in cognate words from Athabaskan languages and dialects. Reconstructed forms are from Rice (1989); cognate sets are from Rice (1989) and Flynn & Fulop (2008).

Pre-form	Slavey	Mountain	Mt Norman	Mt Wrigley	Bearlake	Hare	English
*tts'	tth'a	p'a			kw'a	w'a	'plate'
*tts'	tth'ené	p'ené			kw'ené	w'éné	'bone'
*tts'	tth'ih	p'ih *or* kw'ih		kw'ih	w'ih		'mosquito'
*tts	ʔetthę́		ʔepę́	ʔepę́	ʔekwę́	ʔedee	'caribou' *or* 'meat'
*dz	eníddhę	eníbę			yenígwę	yerígwę *or* yeríbę	'want'
*s	ʔdhé		ʔevéh *or* ʔewéh		wé	ʔwé	'hide, skin'
*s	thę́	fę́			whę́	wę́	'star'
*s	the	fe			whe	we	'belt'
*s	tha	fa			wha	wa	'tent pole'

For Chilcotin, Cook (1981) describes changes of labialized velars into alveolar fricatives or affricates (Examples 5.57–5.58). However, more recent reconstructions propose no labialized obstruents at all for Proto-Athabaskan, instead reconstructing retroflex affricates.

(5.57) *x̯^{h}un > zun 'good'

(5.58) *g̯^{w}e-n > dzin 'day'

Regardless of the complexity within the synchronic and diachronic phonological systems of Athabaskan languages, and the difficulty of establishing the direction of change, there is strong evidence of various types of alternation between labialized velar sounds and dental sounds in Athabaskan languages.

Plain labials to coronals

Developments of labial stops, specifically labial stops with contrastive secondary palatalization and labial stops in front vowel environments, into coronals are rare but observed in a number of language families. Examples are presented in e.g. (Ohala, 1978; Thomason, 1986), some of which are described in this section. Picard (1984, p. 428–429) has argued that these types of changes are typologically impossible and require intermediate stages of development; it is the case that in some instances there are observed intermediate stages

Arapaho-Atsina The Proto-Algonquian labial stop **p* developed into /č/ and /c/ in palatalizing environments in Arapaho-Atsina (Examples 5.59-5.60, from Picard 1984; cf. Example 5.61 in which the outcome is velar before inherited **a*). Picard (1984) proposes an intermediate stage /k/ for these sound changes, but there is no evidence to support this and his argument hinges on the proposed typological unlikelihood of labials developing to palatals; Thomason (1986) argues that an intermediate stage is not necessary given the existence of other examples including those described in this section.

(5.59) Proto-Algonquian *miipiči* to Arapaho *béíči*θ and Atsina *b*^{y}*íícic* 'tooth'.

(5.60) Proto-Algonquian *peešekwani* to Arapaho *čéésé́y* and Atsina *čéé*θ*íy* 'one'.

(5.61) Proto-Algonquian *čiipayaki* to Arapaho θ*iikóno*^{ʔ} and Atsina *ciikóno* 'ghosts'.

Albanian Some dialects of Albanian show palatal reflexes of sequences of labial stops and the palatal approximant /j/ (Thomason 1986, after Desnitskaja 1968).

(5.62) Albanian *pjeshke* to Geg *ćešk* 'peach'.

(5.63) Albanian *plep* to *pjep* to Geg *ćep* 'poplar'.

Bantu Examples of Bantu developments are given in Examples 5.64-5.65, taken from Thomason (1986, p. 184). These and other similar Bantu changes are discussed at length by Ohala (1978), who proposes abrupt changes from [pjʷ] to [tʃʷ] or [tʃ].

(5.64) Proto-Bantu **pia* to Swahili *ša* 'be hot'.

(5.65) Proto-Bantu -pi̯a- 'spit' to Shambaa -*š*, Konde *ṣwa*, Kongo (dial.) -*tswila* and Pedi (Sotho) -*tswa*.

Czech Palatalized Old Czech labial obstruents became coronals in fronting environments and labials elsewhere in some dialects (Bělič, 1966; Thomason, 1986; Andersen, 1973; Ohala, 1978). Examples 5.66-5.69 are taken from Andersen (1973, p. 765); example 5.70 is from Ohala (1978, p. 370). In these articles there are additional examples that demonstrate the change also affected bilabial nasals.

(5.66) Standard Czech /pjɛt/ to East Bohemian /tɛt/ 'five'.

(5.67) Standard Czech /pjiːvɔ/ to East Bohemian /tiːvɔ/ 'beer'.

(5.68) Standard Czech /pjɛknjɛ/ to East Bohemian /tɛknjɛ/ 'nicely'.

(5.69) Standard Czech /prapiːsek/ to East Bohemian /pratiːskɔ/ 'door-post'.

(5.70) Standard Czech *bílý* to East Bohemian /diːlej/ 'white'.

Discussion of the cross-linguistic evidence For the Czech changes described above, Bělič (1966) proposed a progressive articulatory path from labials to coronals; however, Andersen (1973) suggests instead a perceptual reanalysis account in which speakers simplified a three-way phonological contrast—based on the division of a frequency continuum from high tonality (apicals) to intermediate tonality (palatalized labials) to low tonality (labials)—into a two-way contrast. A similar account is proposed in Ohala (1978), in which acoustic similarity between palatalized labials and dentals is credited with prompting phonological reanalysis to explain alternations in Bantu and some Romance languages. This argument is supported by evidence that coarticulatory tongue movement is greater for bilabials than for other stops (Sussman et al., 1991), presumably because tongue position is not a necessary cue to stop place identification for these sounds. Acoustic similarity between palatalized labials and coronals before front vowels has been shown to correlate with perceptual similarity, as in the perceptual confusion experiment described in Winitz et al. (1972) and in the perceptual confusion data presented here in §4.4, in which /pi/ and /ti/ were more likely to be confused than other stops in high front vowel environments. The proposal of a similar analysis for labialized velar changes to palatals is complicated by the fact that acoustic evidence of tongue movement towards the velum may be a necessary cue to place identification for these sounds (although this cue could be provided by other articulatory movements that are equivalent in acoustic output). Whilst the acoustic analysis in Chapter 3 demonstrated that in at least some instances it is possible for the velar articulation to be fronted in front vowel environments, this did not translate to perceptual similarity in Experiment 4. Bateman (2010) argues instead for progressive articulatory accounts of labial palatalization, pointing out that for Tswana changes of /p/ to coronals there are possible intermediate forms with both labiality and coronality, e.g. [tʃ^{w}], [pʃ^{w}], [ps^{w}].

5.3.3 Phonological approaches

Some phonological interpretations provide support for feature specifications which connect labiality to coronality or palatality. Hall-Lew (2006) argues that labio-palatal glides and coronals share the articulatory feature [+ATR] or advanced tongue root, which is associated with a raised and fronted tongue body; this may provide an articulatory motivation for a merger of fronted labialized velars with coronals. Alternatively, explanations may be proposed using acoustic features; some have suggested a link between dentals and sounds with secondary labialization based on the feature [grave]. Howe & Fulop (2005, p. 14ff.) propose that the Athabaskan developments of dentals to labialized velars (as in Table 5.1) can be explained using acoustic features. In Jakobson's original formulation dentals are classified as not [grave] (Jakobson et al., 1951, p. 27–28) on the basis that their description of [grave] stipulates a predominance of energy in one half of the spectrum. Ladefoged (1997) refines this definition, recommending instead that [grave] be defined as the presence of 'salient aperiodic energy in the lower part of the spectrum', which differs from the earlier definition in that it excludes vowels from the class of [grave] sounds. Howe & Fulop (2005) refine the definition further, arguing that the important factor is whether or not low frequency energy is audible, i.e. it is not obscured by high frequency noise; the distinction focusses then on the lack of high frequency noise rather than the presence of low frequency noise. This allows a redefinition of the Athabaskan interdental fricatives as [grave], an acoustic feature which would be intensified by the addition of secondary labialization (which, as discussed in §3.2.5, is flat and so augments grave-ness). Thus the sound change from dental to labialized dental is represented as an intensification of grave-ness by the addition of a [+flat] element. Similar sound changes are observed by Stuart-Smith et al. (2007) for Glaswegian English; stylistic variation between [θ] and [f] is observed, but also occasionally a presumably intermediate variant [θ^w]. If the Greek outcome stops were indeed dental, then a similar account could be proposed for the change of *k^w to /t/. However, there is evidence which indicates that perceptual discrimination between [θ] and [f] is heavily biased: participants in perceptual confusion experiments tend to mistakenly select [f] in response to a [θ] stimulus more often than the reverse (Miller & Nicely, 1955), leading some to consider the change unidirectional (Blevins, 2004; Garrett & Johnson, 2013). The fronting of the velar articulation observed in the British English and Western Zapotec data also complicates this analysis. Vago (1976) uses the feature grave to describe conditioned fronting of [w] in Baule; according to his analysis, [w] becomes [-grave] when it appears between two phonemes classified [-grave]. This is manifested as a front rounded glide which Vago transcribes as [ẅ], but which might be transcribed in IPA orthography as [ɥ] (Pullum & Ladusaw, 1996) (although Pagliuca & Mowrey 1980, p. 504–505 refute this analysis and provide an articulatory account involving coarticulatory tongue fronting). Although the tokens here do not seem to have progressed to that stage (the lowering of F_2 in the following vowel indicates that the tongue position is still further back than for a labiopalatal during the glide), the fronting of the tongue relative to labialized velars in different vowel environments may still affect its classification as [grave]: some British English tokens of /kwi/ had spectra with their dominant peak above 2.5 kHz (Fig. 3.7d), making them [-grave] by any of the definitions mentioned above. Thus the role of [grave] in these Greek sound changes remains rather unclear.

5.3.4 Palatalization

Philologists frequently refer to palatalization in the context of the Greek sound changes $*k^w$, $*g^w$, $*k^{wh}$ to /t, d, tʰ/. In phonetics, palatalization refers to articulations which involve the movement of the tongue towards the hard palate. The definition includes both secondary palatalization, in which the tongue moves towards the palate transitionally between articulations (as in e.g. English *keep*, in which the tongue is more forward at the release of the velar burst than in e.g. English *caught* in preparation for the front vowel [i]) and contrastive palatalization, in which palatalization is a meaningful feature of the stop but occurs secondarily to the primary articulation (e.g. Polish stops /kʲ/ and /gʲ/, contrasting with /k/ and /g/).

Palatalization as a phenomenon encompasses several phonetic characteristics of sounds, all of which need only involve some degree of constriction at the palate. As this is something of a broad term including many potential places and manners of articulation, articulations involving contact of the tongue with the palate are often subdivided into categories. Ladefoged (1982) identifies two varieties of palatalization, primary (in which an articulation "becomes more palatal") and secondary (or the "addition of a high front tongue position, like that of [i], to another articulation"). Keating (1993) uses x-ray tracings and palatograms to identify gradience in the palatalization of plain velars in [i] contexts, prompting a subdivision of palatalization into palatal stops, palatalized velar stops, and fronted velars (Keating & Lahiri, 1993).

Acoustic characteristics of velar palatalization in English, Hungarian, Czech, and Russian are provided in Keating & Lahiri (1993). The authors identify frequency of the spectral peak during the burst portion of the stop as indicative of its degree of palatalization: for example, the peak spectral frequency of back velars in the experiment equalled the F_2 frequency of the following vowel, the peak of the palatal velar equalled or exceeded the F_4 frequency of the following vowel, and the fronted velar peaked somewhere in between. The relationship between articulatory fronting and peak spectral frequency is described as orderly, i.e. the frequency of the spectral peak increases progressively as the tongue is fronted and the size of the front cavity is reduced (Keating & Lahiri, 1993, p. 96). Reduced perceptual distance between fronted velars and coronals is an important component in diachronic velar palatalization (Guion, 1998; Johnson, 2003) suggesting that acoustic variation, interacting with perception, plays a prominent role in these sound changes.

In §3.7 acoustic evidence presented in Chapter 3 was used to support a conclusion that velar stops undergo coarticulatory fronting before the high front vowel /i/ in British English, but the acoustic effects are less strong for the less front vowel /ɪ/. It was concluded that the audibility of the acoustic cue to coarticulatory fronting of labialized velars in front vowel environments depends upon the timing of the secondary articulation, and moreover that this cue differs cross-linguistically and may relate to the phonological status of the sound sequence.

5.3.5 Conclusions concerning developments to coronals

The conclusion in Chapter 3 that coarticulatory fronting of labialized velars can happen, but differs cross-linguistically and depending upon the quality of the following vowel, may account for the relative typological rarity of historical fronting of labialized velars. The spectra for /kwɪ/ and /kwi/ show prominent peaks between F_2 and F_4 of the following vowel, which are not observable in the spectra for /kwɔ/ and which resemble the pattern observed for non-labialized velars with coarticulatory fronting. This indicates that it is possible for articulations of contrastively labialized velars to show acoustic evidence of coarticulatory fronting in front vowel environments. On the one hand, acoustic evidence of tongue fronting may also be overshadowed by the high intensity of noise below 2 kHz for labialized velars, making it an acoustic but not quite an auditory characteristic of the burst. On the other hand, this effect has been shown to vary cross-linguistically; for the British English stops, the upper limit of the confidence interval of the high frequency peak in /kwi/ tokens exceeded that of the low frequency peak (Fig. 3.7d). This suggests that for some tokens the high frequency peak caused by coarticulatory fronting could have some auditory significance. Similarly, the Western Zapotec tokens showed speaker-specific differences, with speaker ZPP-ANT producing /kwi/ bursts with relatively high-intensity high frequency peaks, although they do not exceed the low frequency peaks in intensity (Fig. 3.19f).

The proposed historic labialized velar fronting processes summarized in §5.3.1 all begin with palatalization of the glide, but the results presented here indicate that fronting of the stop burst without palatalization of the secondary articulation is also possible. This suggests that the sound changes observed in Ancient Greek may have begun with fronting of the velar articulation, rather than with any of the sound changes suggested in §5.3.1, which affect the labial secondary articulation first. Thus a third possibility can be proposed, in opposition to the two developmental routes described in §5.3.1:

/kʷ/ > /k̟ʷ/

On the other hand, significant dynamic differences in formant frequencies in the vocoid following labialized and non-labialized velars provide listeners with an acoustic cue to differentiation between /k/ and /kw/; the formant analysis shows a steeply rising second formant transition for /i/ and /ɪ/ after /kw/. This strongly detectable acoustic cue may account for the typological rarity of sound changes which are motivated by coarticulatory fronting of /kw/ in front vowel environments. Reduction or alteration of the secondary articulation may diminish the usefulness of this dynamic cue to /kw/ identification, though; this indicates that the sound change may begin with a change to the secondary articulation, but that change need not necessarily involved palatalization. Different acoustic cues to differentiation between labialized and non-labialized velars are available to listeners depending on vocalic environment, which may account for the conditioning of developments in Ancient Greek according to the quality of the following vowel.

5.4 Developments to velars

5.4.1 Introduction

The loss of phonologically contrastive secondary labialization from labialized stops in positions adjacent to rounded vowels is a typologically common historical linguistic occurrence. The fact that this phenomenon is observable in numerous unrelated languages at different times and locations suggests there is a phonetic basis for this sound change. This chapter provides context for assuming a phonetic basis by describing similar sound changes in a number of different languages and language families, then explores the details of a phonetically motivated account using data and conclusions from the acoustic and perceptual analysis of British English and Western Zapotec presented in Chapters 3 and 4.

5.4.2 Typological parallels: labialized velars to velars

Indo-European languages Delabialization of labialized velars adjacent to rounded vowels in Indo-European, Greek, and other Indo-European daughter languages was discussed in §2.6.

Uto-Aztecan languages There is strong evidence that delabialization of labialized velars in back rounded vowel environments took place in Proto-Uto-Aztecan. If delabialization of labialized velars adjacent to rounded vowels had taken place in Proto-Uto-Aztecan, we would expect to see a higher proportion of roots in which a rounded vowel appears adjacent to a velar, due to the falling together of the velars and the labialized velars in this position. This is exactly what is observed. The inventory of Proto-Uto-Aztecan roots reconstructed by Stubbs (2011) shows an unexpectedly high number of root-initial syllables beginning with the velar stop **k* followed by rounded vowels **o* and **u*. The combined number of roots with **ko* or **ku* is 65, notably higher than roots with other stops in these vocalic contexts (e.g. **po/pu* at 49 or **to/tu* at 50) and the overall average of 30 Stubbs (2011, p. 13). Perhaps relatedly, there are only two reconstructed roots with **w* before **o* or **u*.

Some Uto-Aztecan languages show a difference in reflexes of labialized velars dependent upon the following vowel. Reflexes of inherited *k^w in O'odham and Nahuatl differ depending upon the following vowel, with *$k^w a$ giving /kʷa/, whilst *$k^w o$ or *$k^w u$ gives /ko/ (Stubbs, 1995).

(5.71) Od. *-ko* 'in, on, at a place' and N. *-ko* 'locative; in, among', cf. Yaqui and Mayo *-bo/-po* 'at a place'.

(5.72) Od. *hikul* 'plant' from proto-Od. *$sik^w ul$ by metathesis from U.-A. *$k^w isul$ 'plant'.

(5.73) N. *-ko* 'locative; in, among' (retains k^w elsewhere).

Evidence of lexical diffusion of a sound change may survive in the Eudeve language, which with its close relative Opata forms the Opatan subfamily of Uto-Aztecan (Shaul, 2010). In Eudeve, inherited Uto-Aztecan *k^w, in positions preceding a rounded vowel, sometimes develops into a bilabial but at other times appears to have become delabialized; extensive evidence of labial developments of *k^w before rounded vowels are presented in §5.2.2, Examples 5.5–5.12. Examples of velar development of *k^w before rounded vowels are presented in 5.74–5.75.

(5.74) Possibly Eu. *kóranan* 'untie', from proposed *k^w*u(ʔ)ta* 'untie', cf. Yq. *búta* 'untie'.

(5.75) Eu. *kowát* 'point' from proposed *k^w*awi* 'point, edge, sharpen', cf. Yq./My. b^w*awí-te* 'sharpen', N. k^w*awi* 'tree, stick'.

Athabaskan languages The Slave languages of northwest Canada show free variation between /kwo/ and /ko/ for many speakers:

(5.76) *dekwo* or *deko* 'it is yellow', cf. *deko* 's/he coughs' (Rice, 1989, p. 33).

Oceanic languages Some Oceanic languages show evidence of delabialization of an inherited proto-Oceanic labialized velar adjacent to a rounded vowel. In examples 5.77-5.79 (Ross, 2011) labialization is reinterpreted as a contrastive attribute of the following vowel rather than of the labialized stop.

(5.77) Mbula *koro* 'shark' from POc. *k^w*arawa* 'shark'.

(5.78) Malasanga *kuro* 'shark' from POc. *k^w*arawa* 'shark'.

(5.79) Proto-Southeast-Solomonic **koi* 'say, talk' from POc. *k^w*ai* 'say'.

Hong Kong Cantonese Cantonese spoken in Hong Kong has both aspirated and unaspirated labialized velar stops but in the latter half of the 20th century /k/, /$\mathrm{k^h}$/ and /$\mathrm{k^w}$/, /$\mathrm{k^{hw}}$/ have merged to give /k/ and /$\mathrm{k^h}$/ in positions preceding syllable rhymes /ɔː/, /ɔːŋ/ and /ɔːk/ (Bauer & Benedict, 1997; Gui, 2005) as in Example 5.80.

(5.80) Hong Kong /kɔːŋ/ for /kwɔːŋ/ 'bright' (Bauer & Benedict, 1997, p. 543)

However, the labialized velar is still produced in careful speech, and there is some evidence that this change may be reversing under the influence of Putonghua following the return of Hong Kong to China (Bauer & Benedict, 1997, p. 336).

5.4.3 Coarticulatory motivations

Phonologization of coarticulatory labialization Phonologization of labialization resulting from coarticulation with a rounded vowel, effectively the reverse of the sound change under discussion, is also attested. Bidirectionality supports a perception-based account, to a certain degree; if two sounds or sound sequences are sufficiently similar to be misparsed as one another, it is expected that generalization in either direction could take place. However, it is not necessary that generalization take place, and this is not the case for all instances of phonetically motivated sound change: for example Guion (1996) finds that the perceptual tendency to hear /ki/ as /tʃi/ was unidirectional. Plauché et al. 1997 relate this to the presence or absence of specific acoustic cues, suggesting that listeners may miss a cue that is present but are unlikely to mistakenly introduce a cue that is not present in the signal. Lynch (2002) outlines historical changes in Proto-Oceanic in which stops adjacent to rounded vowels, and therefore subject to non-contrastive coarticulatory rounding, became contrastively rounded in some Oceanic daughter branches through a process of phonological reanalysis.

Perception of labialized velars in back round vowel environments Data from British English show that mean formant frequencies of vocoids in the /kwo/ category overlapped with those in the /ko/ category and were in fact statistically indistinguishable at most time points throughout the vocoid; see Chapter 3 for a complete discussion of the data. This acoustic similarity was reflected in perceptual similarity: in Experiment 3, perceptual distances of 0.65 (male speakers) and 0.45 (female speakers) were recorded for /kwo/ and /ko/, making them the fourth (males) and second (females) most similar pair of CV sequences tested. However, this raises the question of why sounds that rank higher on this scale of perceptual similarity than /kwo/ and /ko/ do not also undergo reanalysis; the case of /kwo/ and /ko/ may be influenced by articulatory similarity as well as perceptual similarity, unlike for example /po/ and /ko/, and /to/ and /ko/, which rank higher in perceptual similarity but involve different articulators.

Acoustic explanations for sound change The acoustic analysis in Chapter 3 revealed similarities in formant transitions between /kw/ and /k/ in back round vowel environments in British English. This suggests there may be a motivation for a 'hypercorrection' style sound change (Ohala, 1993). Under this model, listeners would reinterpret the acoustic cues to lip rounding as a characteristic of the vowel rather than as a characteristic of the stop. Working on the assumption that the rounding of the stop is due to coarticulation under the influence of the vowel, listeners may infer that the stop itself is not contrastively labialized. This account depends upon the sequences sounding sufficiently similar to one another in this vocalic environment that phonological reanalysis along these lines is possible. There is some evidence indicating that speaker-listeners can interpret similar phonetic input as multiple different phonological sequences depending upon the morphological or syntactic context in which it appears (Hale, 2012); this implies that it may be possible for a phonological reanalysis to take place without a phonetic change having taken place first.

The hypercorrection hypothesis suggests a degree of coarticulation between the stop and following vowel, providing a context for reinterpretation. There is evidence that the tongue body shows greater context-dependent variability in producing back vowels than front vowels (Recasens, 1999; Perkell & Cohen, 1989; Stevens & House, 1963). Recasens (1985) found that, in Catalan, F_3 frequency was more variable for velars in different vowel environments than for other stops. Velar stops, rounded vowels, and labial-velar approximants are all characterized by low F_2 frequencies (Recasens, 1985; Ohala & Lorentz, 1977). Perkell & Matthies (1992) found that the onset of lip protrusion for stops preceding rounded vowels began earlier when the duration of the consonant was longer; although I could not gather data for closure durations because the stops were utterance-initial, velar stops had the longest voice onset times (defined in that study as the interval from the burst to the onset of voicing) in the laboratory recorded data in Chapter 3, indicating a greater potential for early-onset coarticulatory lip rounding. This suggests that there will be a high degree of coarticulation between velars and following rounded vowels.

Quantifying degree of coarticulation using locus equations

Introduction The degree of coarticulation between stops and following vowels can be quantified partly by deriving locus equations from the formant frequencies of the transition and vowel. Locus equations represent regression lines which model the relationship between vowel formant frequencies at the start of the vocoid (i.e. the point at which coarticulatory effects of a preceding stop will be manifested most prominently) versus the temporal midpoint of the vowel (by which time the articulators are supposedly closest to their 'target' positions for that vowel) (Lindblom, 1963a; Sussman et al., 1991; Sussman & Shore, 1996; Sussman, 2015; Iskarous et al., 2010) and thus can be used to infer the place of articulation during the stop closure (Krull, 1988, p. 66). The equation shows in effect the extent to which vowels are modified from their target by a specific preceding stop, although it should be noted that gestural undershoot may extend the effects of coarticulation throughout the duration of the vocoid, including its mid-point 'target'; see formant transitions in Fig. 3.1d and 3.2d, in which F_2 frequencies for /ɪ/ and /i/ are lower after /kw/ than after other stops throughout the production of the vowels. Locus equations are expressed as in equation (5.81) (adapted from Krull (1988, p. 67)), where F_{nf} denotes the locus of formant frequency F_n, F_{nt} denotes the target frequency of F_n, and k_f and c_f are coefficients which vary by stop place and vowel quality.

$$F_{nf} = c_f + k_f F_{nt} \tag{5.81}$$

The constant k represents the slope of the regression line, which varies for F_2 depending on the place of articulation of the stop (Lindblom, 1963a,b). Sussman et al. (1991) and Sussman & Shore (1996) claim that locus equations can be used to identify stop place articulation and may function as perceptual cues for listeners; Fowler (1994) disputes this claim after finding that changes in manner of articulation altered the coarticulatory relationship between the stop and the vowel in CV sequences, meaning the locus equations were not useful as cues to identifying place. However, that debate is not consequential for the current investigation in which locus equations are used to define degree of coarticulation.

The question of how much and what kinds of information locus equations can provide about coarticulation of adjacent stops and vowels has been disputed. Allen et al. (1987, p. 113–115) found locus theory to be useful in modelling /C/ to /V/ transitions in speech synthesis. Krull (1988, p. 66–68) suggests that locus equations could be used to describe the degree of coarticulation between stops and following vowels. According to this theory, when the locus frequency varies as a direct and linear function of the vowel target frequency, i.e. there is a maximal coarticulation effect, then $k = 1$. When there is no coarticulation effect, i.e. the locus frequency remains constant in different vowel environments, $k = 0$ (Krull, 1988; Lindblom & Sussman, 2012). This concept was challenged by Löfqvist (1999), who found that regression slopes did not have any significant relationship with measures of coarticulation provided by sensors recording the movement of the tongue and lips. Tabain (2000, 2002) found that locus equations accurately reflected the degree of coarticulation observed using electropalatography for voiced stops, but did so less successfully for voiceless stops and for fricatives; the lack of accuracy for voiceless stops could be due to the articulators moving towards the vocalic target during the aspiration period before the first formant measurement is taken. However, when the degree of coarticulation was controlled using articulatory simulations with differing intergestural timing, locus equations did accurately reflect the degree of coarticulation Chennoukh et al. (1997). Similarly, Lindblom (1998) used simulations of different amounts of coarticulation to demonstrate a relationship between degree of coarticulation and locus equation slopes.

Locus equations have only occasionally been derived for stops with contrastive secondary labialization: Denzer-King (2013) found steeper slopes for non-labialized velars than for labialized velars, indicating a greater degree of coarticulation for the former than for the latter, although no contrastively rounded vowels were studied. In positions before rounded vowels one would expect locus equations to indicate that both contrastively labialized and contrastively non-labialized velar stops show coarticulation effects towards the following rounded vowel.

In the case of contrastively labialized velars preceding rounded vowels, lip-rounding at the burst is expected due to the contrastive labialization inherent to the stop. Due to the presence of lip-rounding and tongue-backing throughout the utterance, the slope coefficient is expected to have a value close to 1. The question then is whether or not, for the /kɔ/ tokens, coarticulatory rounding of the stop under the influence of /ɔ/ is sufficient to render the relationship of formant frequency at the onset and at the vowel target similar to that for /kwɔ/ tokens.

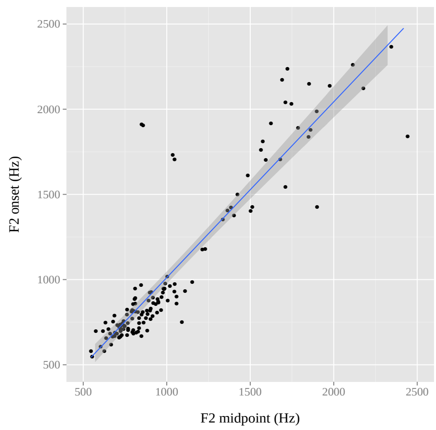
Methods Locus equations were derived by plotting the F_2 and F_3 frequencies at voice onset against the respective F_2 and F_3 frequencies at the vowel target for each stop, determined here to be after 0.05 s of voicing (by which time it is expected the effects of stop coarticulation on the vowel are most reduced) and calculating a regression line using the `lm()` function in the R programming language.

Two statistical outliers were removed from the dataset for /kɔ/ sequences. Means and standard deviations were calculated for the formant frequency measurements at the onset and at 0.05 s from the onset of the vocoid; data points with values under or over three standard deviations from the mean were discarded. In the case of /kɔ/ sequences, mean F_2 frequency at vocoid onset was 1055 Hz and the standard deviation was 362 Hz; a datapoint with an F_2 frequency value of 2550 Hz was discarded. At the measurement time, mean F_2 frequency was 970 Hz and the standard deviation was 383 Hz, so a datapoint with an F_2 of 2205 Hz was discarded.

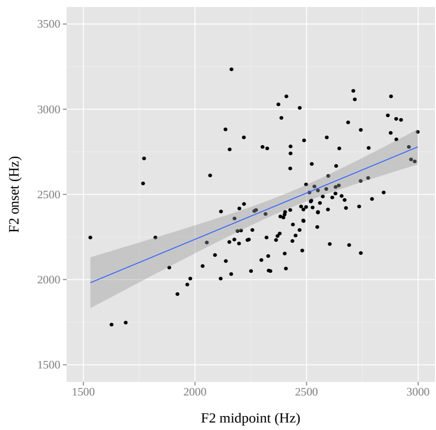
Results Locus equations for *quarter* and *caught* words are compared below.

Fig. 5.1: Locus equations for F_2 and F_3 of *caught* and *quarter* words.

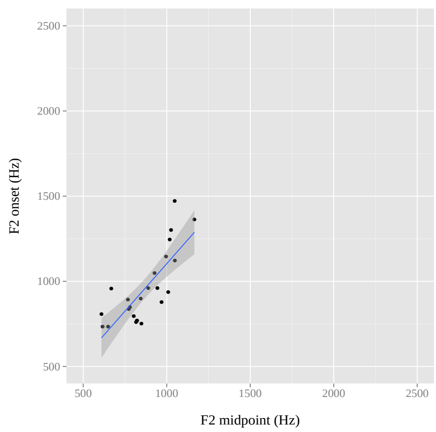
(a) *Quarter* F_2; $n = 132$; $y = 202 + 0.80x$; $r^2 = 0.83$.



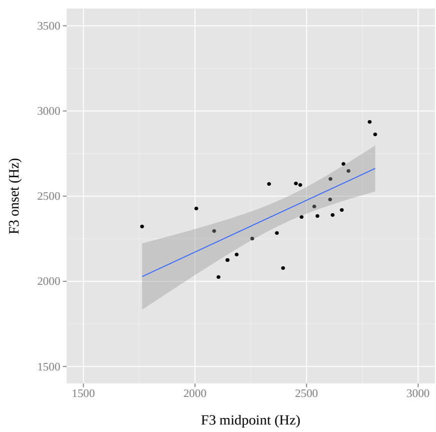
(b) *Quarter* F_3; $n = 132$; $y = 783 + 0.68x$; $r^2 = 0.41$.



(c) *Caught* F_2; $n = 23$; $y = 342 + 0.55x$; $r^2 = 0.60$.



(d) *Caught* F_3; $n = 23$; $y = 499 + 0.79x$; $r^2 = 0.46$.



Discussion The slope for F_2 of /kwɔ/ sequences is steep with a coefficient of 0.8. This is indicative of a high degree of coarticulation according to the theories outlined by Krull (1988) and Lindblom & Sussman (2012) and confirms the prediction made in the Introduction to this section (Fig. 5.1a). For F_2 of /kɔ/ sequences the coefficient is 0.55 (Fig. 5.1c), suggesting some coarticulation; however, the coarticulatory effect is less than is apparently observed for the /kwɔ/ tokens. In the F_3 space, there is a higher coefficient for /kɔ/ than for /kwɔ/ but the variance is greater than for the F_2 data indicating a weaker correlation.

Sussman et al. (1991) and Sussman & Shore (1996) observed that the locus equations for American English /g/ before back vowels and for /b/ overlapped and that their y-intercept values were not significantly different; the r^2 values observed here are more similar to those described for English labial stops by Sussman et al. (1991). Likewise, Sussman et al. (1991) and Sussman & Shore (1996) observed y-intercepts in the range 384-1009 Hz for F_2 frequencies after English velars, and similarly high y-intercepts for Urdu, Cairene Arabic, and Thai. In the data shown here, both the y-intercept for the F_2 frequency of /kɔ/ at 342 Hz and the y-intercept for /kwɔ/ at 202 Hz are below the lower limit of the range of Sussman's data for velar stops. Even when compared to Sussman's measurements of allophonic variants of velar stops in back vowel environments the y-intercepts are still lower; Sussman found a mean y-intercept of 559 Hz for velarized /g/, much lower than for palatalized /g/ but not as low as for labials. However, the y-intercept for /kwɔ/ is within the range observed by Sussman for bilabial stops (69–257 Hz), whilst the y-intercept for /kɔ/ is higher; this suggests a difference in labialization effects between /kwɔ/ and /kɔ/.

Whilst in the F_2 space the /kwɔ/ tokens have a higher slope coefficient, this is counterbalanced by the higher y-intercept of the /kɔ/ tokens. For example, for a vowel with an F_2 target of 700 Hz (the F_2 example frequency provided for the vowel /ɔ/ by Catford 1977b) the projected locus for /kwɔ/ is 762 Hz; for the /kɔ/ tokens the projected locus is 727 Hz, a difference of only 35 Hz. In the F_3 space, the reverse is true: /kwɔ/ tokens have a shallower slope and a higher y-intercept, but with an F_3 target of 2470 Hz (the mean of all /ɔ/ F_3 frequencies measured in the BNC data in Chapter 3) the /kwɔ/ locus is 2463 Hz, compared to 2450 Hz for /kɔ/. Thus for a given target F_2 and F_3 frequencies, /kwɔ/ tokens will have slightly higher frequency onsets than /kɔ/ tokens; but the difference is small and may not be perceptible to listeners. In §3.1.1 the JND (Just Noticeable Difference) for vowel formant frequencies was estimated at 1–5%; the differences in projected loci for /kwɔ/ and /kɔ/ are within 1% of the formant frequencies for F_3 (7 Hz difference) and within 5% for F_2 (35 Hz difference), suggesting that for many tokens the difference may not be auditorily perceptible. In other words, [kʷ] in /kw/ ≈ [kʷ] in /kɔ/.

5.4.4 Conclusions concerning developments to velars

The review of cross-linguistic evidence presented in §5.4.1 indicates a phonetic motivation for the sound changes, on the basis that similar changes occurring frequently in unrelated and geographically distant languages indicate a physiological motivation rather than a language-specific or phonological motivation. It was not possible to conduct acoustic analysis of Western Zapotec labialized velars preceding labialized vowels because labialized velars do not appear in that position in Zapotec (this fact in itself could be considered further evidence for a cross-linguistic tendency towards loss of contrastiveness of labialized velars in this position). The acoustic study described in Chapter 3, and expanded here with the analysis of locus equations, suggests that British English /kwɔ/ and /kɔ/ are acoustically similar. The locus equation analysis indicates that coarticulation of a non-contrastively labialized velar under the influence of a following rounded vowel results in a formant trajectory that is very similar to that of a contrastively labialized velar in the same vocalic environment. Coarticulatory effects of a following back rounded vowel lead to similar locus equations for each sequence, although a lower *y*-intercept for contrastively labialized velars also renders them similar to bilabials. This acoustic similarity translates directly to perceptual similarity, as demonstrated by Experiment 3 in Chapter 4. However, high levels of perceptual confusion between e.g. /ko/ and /to/, which do not seem to show a similar cross-linguistic tendency to merge phonologically, suggest that articulation works alongside perception in motivating this sound change. Although exact realizations of labialized velars in other languages are likely to differ, this analysis provides a template by which acoustic and articulatory similarity caused by coarticulation of velar stops adjacent to rounded vowels could lead to scenarios in which some variant productions of contrastively labialized velars in back round vowel environments are perceptually indistinguishable from productions of non-contrastively labialized velars. This provides a motivation for phonological reanalysis of contrastively labialized velars as non-contrastively labialized velars in this environment.

A situation was observed in Ancient Greek in which some labialized velars become coronals before mid front vowels but not before mid high vowels (Chapter 2). This seems contrary to the examples from other languages in which velars became coronals more readily before high front vowels than before less high front vowels (Guion, 1996, 1998). However, in the acoustic data in Chapter 3, it was observed that before British English /i/ there was a highly distinctive formant transition in which F_2 rose sharply following the release of the velar constriction. This potential cue was less prominent for /ɪ/. Thus a cue to velar fronting before /ɪ/, whilst less prominent than for /i/, may be less susceptible to being obscured by low frequency noise from the following labial-velar glide and thus more audible. This may provide some explanation for why a labialized velar sound may be more likely to palatalize before a mid front vowel than before a high front vowel, relative to a non-labialized velar; the Greek vowel was /e/, which was presumably more front than /ɪ/ but less close than /i/, which may affect how these two factors (fronting at the burst and rising of F_2 frequency in transition) interact with one another.

6
Conclusions

6.1 Introduction

This thesis set out to investigate the mechanism of contextually conditioned sound changes of labialized velars to labials, coronals, and non-labialized velars in Ancient Greek. Previous studies of the sound changes had suggested some specific processes by which the conditioned outcomes may have come about:

(a) **Labial outcomes**

Ohala (1990b, 1993) suggested a perceptual confusion account in which labialized velars became labials through perceptual misparsing, based on a supposed acoustic similarity between labialized velars and labials. This was proposed to explain sound changes in Greek, other Indo-European languages, and many unrelated languages which showed similar historical sound changes. Whatmough (1937) and Garrett & Johnson (2013) describe an alternative scenario in which incremental changes lead to intermediate double articulation labial-velar stops before velarity is lost.

(b) **Coronal outcomes**

Philologists have proposed articulatorily incremental developments from PIE labialized velars to Greek coronals, based on the attestation of apparently intermediate forms in Arcadian inscriptions and glosses (Lejeune, 1972; Risch, 1979; Brixhe, 1996; Bubeník, 1983) and in Lycian inscriptions (Melchert, 1994, 2004). Most of these explanations involved progressive palatalization of the secondary glide, i.e. developments from $*k^{w}$ $*g^{w}$, $*g^{wh}$ to /$k^{ɥ}$/, /$g^{ɥ}$/, /$g^{ɥh}$/.

(c) **Velar outcomes**

This sound change has not been exposed to much rigorous analysis, perhaps due to its rather intuitive nature. Articulatory and acoustic similarity in the stop and vowel transition is generally supposed to have resulted in merger either in Greek or in Proto-Indo-European (Weiss, 1994).

Chapters 3 and 4 sought to refine our understanding of labialized velars by investigating their acoustic properties in British English and Western Zapotec, and by identifying perceptual biases in British English. The focus was on describing what might be termed 'microvariations', or small coarticulatory and dynamic differences in the realization of labialized velars depending upon their context. Chapter 5 brought together the acoustic and perceptual evidence to assess the plausibility of each proposed mechanism for the sound changes.

This concluding chapter ties together the findings made in each part of the thesis to provide, to the extent possible, a plausible account of each sound change.

6.2 Empirical findings

The acoustic experiments in Chapter 3 and the perception experiment in Chapter 4 provided several major findings for each of the vocalic contexts studied. They are presented in this section with reference to each vocalic context for the acoustic data, and synoptically for the perception data.

6.2.1 Acoustic experiments

Labialized velars in all vocalic contexts Regardless of the following vowel, the burst spectra of all the labialized velars had a prominent peak at roughly the equivalent of F_1 of the following vowel. This confirms their classification as [+grave] in the Jakobsonian acoustic feature system.

Labialized velars before front vowels Transitions before front vowels were characterized by a steeply rising F_2 trajectory. For British English, this was steeper before /i/ than before /ɪ/; for Western Zapotec, the F_2 rise was less steep than for British English and the lowering effect less strong. F_3 frequency was lower for all front vowels after labialized velars than after /p, t, k/ in British English. In Western Zapotec, F_3 was not consistently lowered by secondary labialization; in fact, two speakers had higher F_3 frequencies after labialized velars. British English /kwi/ had a prominent high frequency peak at the burst that was not present in other vocalic contexts and indicated coarticulatory fronting of the velar articulation. Although phonologically these sounds are still [+grave] within British English, given the lack of a contrasting palatal sound, the parameters for phonetic classification of [grave] may no longer be met due to the effects of coarticulatory fronting.

Labialized velars before back vowels Before Western Zapotec /a/, labialized velars generally had a lowering effect on F_2 and little effect on F_3 relative to non-labialized velars. Before British English /o/, labialized velars had a lowering effect on F_2 and F_3 relative to other stops tested. F_3 was a more prominent acoustic cue to differentiation between /kwo/ and /po/.

Labialized velars before back round vowels The formant transitions of round vowels after contrastively and non-contrastively labialized velars in British English differed significantly in F_2 frequency, but only by a margin of 68 Hz. The bursts, on the other hand, differed more; the /kwo/ burst was relatively diffuse, with a more even concentration of energy across the frequency range, whereas the /ko/ was compact and dominated by a low frequency peak.

6.2.2 Perception experiment

Participants heard tokens consisting of monosyllabic English words with the stops /p, t, k, kw/ as onsets and the vowels /e, i, o/ as nuclei. Stimuli were obscured by background conversational noise. Participants correctly identified /kw/ before front vowels in nearly all trials, and were not more likely to identify /kw/ as /p/ in any of the vocalic environments included in the perception task. There was, however, a strong tendency to identify /kwo/ as /ko/ and vice versa.

6.3 Theoretical implications

Neither the acoustic nor the perception experiments provide a basis for supposing that an immediate sound change of labialized velars to labials is plausible for Ancient Greek or for other languages in which this sound change is attested. Evidence from a number of African and North American languages surveyed in Chapter 5 shows that labialized velars alternate—either dialectally or as allophonic variants—with labial-velars, and that labial-velars show developments to labials via gradual lenition of the velar articulation. Whilst there is no evidence of this taking place in the Indo-European languages in which labial outcomes are attested, there is also no evidence of it not taking place, due to the long periods without literary records and the opacity of writing systems such as Linear B.

In Chapter 2, a pattern was observed in which labialized velars appear to be more likely to show coronal outcomes before mid front vowel /e/ than before high front vowel /i/. This has been considered unusual on the basis that, in many languages, palatalization of velars before /i/ is a prerequisite to palatalization of velars before /e/ (Stephens & Woodard, 1986). However, the acoustic study in Chapter 3 shows that it is unwise to consider labialized velars equivalent to velars in this respect. It was found that, for British English, there was acoustic evidence of coarticulatory fronting of labialized velars before front vowels; this was was more prominent the more front the vowel. However, the dynamic cue to labiovelarization of the vowel was also more prominent the more front the vowel, obscuring the cue to velar fronting. This somewhat paradoxical relationship emphasises the importance of relative timing of gestures when discussing sound changes involving stops with multiple places of articulation.

Timing was also found to be important in discussing cross-linguistic attributes of labialized velar sounds. For British English, a sequence /kwi/ might generally be transcribed as [k^wwi], whilst for Zapotec it might be transcribed [k^wi], with the articulations reflecting the difference in phonological status of the sound sequence in each language. However, the realisation at the burst was actually different for each language, with Zapotec showing much more acoustic evidence of labialization and much less vowel-conditioned variation at the burst than British English. For British English, the vowel formant transitions showed much more evidence of labiovelarization than the burst, which was then more susceptible to showing auditorily significant contextual variation. Thus it seems that for contextually conditioned variation to become embedded as historical change in the primary articulation, it may be a prerequisite for there to exist a timing mismatch with the secondary articulation, possibly related to phonological reanalysis.

Timing was equally relevant to the differentiation between contrastively labialized and non-labialized velars in positions preceding back round vowels. There was relatively little difference in formant transitions for these sequences after the initial measurement, with differences encoded more in the burst spectra; /kwo/ tokens had a less compact distribution of energy and more energy above 3 kHz than /ko/ spectra, although both had dominant peaks equivalent to F_1 and F_2 of the following vowel. Locus equations in Chapter 5 investigated whether the coarticulation between /k/ and /ɔ/ was sufficient to render its locus more similar to that of /kwɔ/; it was found that the projected loci for /kwɔ/ and /kɔ/ were close enough that the difference was on the threshold of qualifying as a JND (Just Noticeable Difference, see §3.1.1). This was sufficient to pose difficulties in listeners' ability to distinguish between the two obstruents in the perception task described in Chapter 4. Thus speakers retained a distinction between the two sound sequences but listeners were unable to successfully recover this distinction. The combination of acoustic and articulatory similarity observed for these sounds recalls the model proposed by Bybee (2010) and discussed in §1.4, in which articulatory and acoustic tendencies within the population work in tandem to produce sound changes.

To conclude, this thesis has presented new and empirically supported analyses of three historical sound changes. The research presented here provided empirical answers to a problem which had been discussed theoretically, namely whether or not labialized velar to labial sound changes have a basis in acoustic misperception. The answer is yes, probably, but perhaps not in the way we thought. Acoustic misperception may contribute to a sound change of labialized velars to labial-velars, due to an acoustic similarity in the effects on formant transitions. However, the extent of acoustic and perceptual similarity between these stops has not been exhaustively investigated. The findings relating to labialized velars before front vowels are new and worthy of further attention. Whilst for the labial outcomes this thesis began with a suggestion by Ohala (1989, 1993) that the sound changes were primarily auditory and ended with the conclusion that they may have more to do with articulation, for the coronal outcomes the reverse is the case: after taking as a starting point the philological assumption of gradual palatalization of the glide, evidence was presented that there is an auditory cue to coarticulatory fronting before front vowels. This cue might in fact be more prominent before less front vowels than before more front vowels, a finding which would benefit from closer investigation, particularly with reference to the Greek sounds; the conditioning vowel for the Greek changes was /e/, which is more front than /ɪ/ and less close than /i/. Both the acoustic and perceptual studies demonstrate how sounds which are acoustically and articulatorily similar may merge. In the case of labialized velars before back round vowels, the coarticulation of the non-contrastively labialized velar to the round vowel was sufficient to render its locus extremely similar to that of the contrastively labialized velar. Variation about the mean indicates that for some tokens these sounds are distinguishable; thus there is a situation where distinguishable and indistinguishable variants co-exists across speakers or utterances, indicating that a future statistical likelihood to favour indistinguishable variants may result in effective merger.

Bibliography

Adams, D. Q. (2014). *Etymological dictionary of Tocharian B*. Leiden: Brill. URL www.brillonline.com

Alderete, J. D., & Frisch, S. A. (2007). *The Cambridge handbook of phonology*, chap. Dissimilation in grammar and the lexicon, (pp. 379–398). Cambridge: Cambridge University Press.

Ali, A. M. A., Van der Spiegel, J., & Mueller, P. (2001). Acoustic-phonetic features for the automatic classification of stop consonants. *IEEE Transactions on Speech and Audio Processing*, *9*(8), pp. 833–841.

Allen, J., Hunnicutt, M. S., Klatt, D. H., Armstrong, R. C., & Pisoni, D. B. (1987). *From text to speech: The MITalk system*. Cambridge: Cambridge University Press.

Allen, W. S. (1957). Some problems of palatalization in Greek. *Lingua*, *7*, 113–133.

Allen, W. S. (1987). *Vox Graeca: A guide to the pronunciation of classical Greek*. Cambridge: Cambridge University Press.

Allen, W. S. (1989). *Vox Latina: A guide to the pronunciation of classical Latin*. Cambridge: Cambridge University Press.

Andersen, H. (1973). Abductive and deductive change. *Language*, *49*, 765–793.

Aura Jorro, F. (1993). *Diccionario micénico*. Madrid: Consejo superior de investigaciones científicas.

Bagwell, C. (1998). SoX – Sound eXchange. URL http://sox.sourceforge.net/sox.html

Bartoněk, A. (2003). *Handbuch des mykenischen Griechisch*. Heidelberg: Winter.

Bateman, N. (2010). The change from labial to palatal as glide hardening. *Linguistic Typology*, *14*(2-3), 167–211.

Bauer, R. S., & Benedict, P. K. (1997). *Modern Cantonese Phonology*. Trends in linguistics: Studies and monographs, 102. Berlin: Mouton de Gruyter.

Becker, K. (2009). /r/and the construction of place identity on New York City's Lower East Side. *Journal of Sociolinguistics*, *13*(5), 634–658.

Beckman, M., Jong, K. D., Jun, S.-A., & Lee, S.-H. (1992). The interaction of coarticulation and prosody in sound change. *Language and Speech*, *35*(1, 2), 45–58.

Beddor, P. S. (2009). A coarticulatory path to sound change. *Language*, *86*(4), 785–821.

Beddor, P. S. (2012). Perception grammars and sound change. In M.-J. Solé, & D. Recasens (Eds.) *The Initiation of Sound Change*, (pp. 37–56). Amsterdam: John Benjamins Publishing.

Beekes, R. S. P. (1995). Aithiopes. *Glotta*, *73*(1), 12–34.

Beekes, R. S. P. (2002). Hom. *géphura*, and Arm. *kamurǰ* 'bridge'. *Glotta*, *78*(1), 12–21.

Beekes, R. S. P. (2010). *Etymological dictionary of Greek*. Leiden: Brill. URL www.brillonline.com

Bell, A. (2002). Back in style: Reworking audience design. In P. Eckert, & J. Rickford (Eds.) *Style and Sociolinguistic Variation*, (pp. 139–169). Cambridge: Cambridge University Press.

Blench, R. (2004). The Benue-Congo languages: a proposed internal classification. URL http://rogerblench.info/Language/Niger-Congo/BC/General/Benue-Congo%20classification%20latest.pdf

Blevins, J. (2004). *Evolutionary Phonology*. Cambridge: Cambridge University Press.

Blumstein, S. E., & Stevens, K. N. (1979). Acoustic invariance in speech production: Evidence from measurements of the spectral characteristics of stop consonants. *The Journal of the Acoustical Society of America*, *66*(4), 1001–1017.

Blust, R. (1990). Three recurrent changes in Oceanic languages. In J. Davidson (Ed.) *Pacific island languages: Essays in honour of GB Milner*, (pp. 7–28). London: University of London, School of Oriental and African Studies.

Blust, R. (2004). *t to k: An Austronesian sound change revisited. *Oceanic Linguistics*, *43*(2), 365–410.

Booker, K. M. (1993). More on the development of Proto-Muskogean *kw. *International Journal of American Linguistics*, *59*(4), 405–415.

Bouchard, K. E., & Chang, E. F. (2014). Control of spoken vowel acoustics and the influence of phonetic context in human speech sensorimotor cortex. *The Journal of Neuroscience*, *34*(38), 12662–12677.

Bouchard, K. E., Mesgarani, N., Johnson, K., & Chang, E. F. (2013). Functional organization of human sensorimotor cortex for speech articulation. *Nature*, *495*(7441), 327–332.

Brixhe, C. (1996). *Phonétique et phonologie du grec ancien*. Louvain-la-Neuve: Peeters.

Bronkhorst, A. W. (2000). The cocktail party phenomenon: A review of research on speech intelligibility in multiple-talker conditions. *Acta Acustica*, *86*(1), pp. 117–128.

Browman, C. P., & Goldstein, L. (1990). Tiers in articulatory phonology, with some implications for casual speech. In J. Kingston, & M. Beckman (Eds.) *Papers in laboratory phonology I: Between the grammar and physics of speech*, (pp. 341–376). Cambirdge: Cambridge University Press.

Browman, C. P., & Goldstein, L. (1992). Articulatory phonology: An overview. *Phonetica*, *49*(3-4), 155–180.

Bubeník, V. (1983). *The phonological interpretation of Ancient Greek: A pandialectal analysis.* Toronto: University of Toronto Press.

Buck, C. D. (1955). *The Greek dialects.* Chicago: University of Chicago Press.

Bělič, J. (1966). K otázce zaniklého litomyšlského t, d, n za někdejší p', b', m'. *Slavica Pragensia*, (8), 1–3.

Bybee, J. (2001). *Phonology and language use.* Cambridge: Cambridge University Press.

Bybee, J. (2012). Patterns of lexical diffusion and articulatory motivation for sound change. In M.-J. Solé, & D. Recasens (Eds.) *The initiation of sound change: Perception, production and social factors*, (pp. 211–234). Amsterdam: John Benjamins Publishing.

Bybee, J. L. (2010). *Language, usage and cognition.* Cambridge: Cambridge University Press.

Bye, P. (2011). Dissimilation. In M. Van Oostendorp, C. J. Ewen, E. Hume, & K. Rice (Eds.) *The Blackwell Companion to Phonology.* Oxford: Wiley Blackwell.

Cahill, M. (1999). Aspects of the phonology of labial-velar stops. *Studies in African Linguistics*, *28*(2), 155–184.

Capo, H. B. (1991). *A comparative phonology of Gbe.* Berlin: Foris.

Catford, J. C. (1977a). Diachronic phonetics. Department of Linguistics, University of Michigan ('originally intended to be Chapters 13 and 14 of Catford Fundamental Problems in Phonetics, 1977').

Catford, J. C. (1977b). *Fundamental problems in phonetics.* Edinburgh: Edinburgh University Press.

Catford, J. C. (1977c). Mountain of tongues: The languages of the Caucasus. *Annual Review of Anthropology*, *6*, 283–314.

Catford, J. C. (1988). *A practical introduction to phonetics.* Clarendon Press Oxford.

Chadwick, J. (1988). Differences and similarities between Cypriot and the other Greek dialects. In J. Karageorghis, & O. Masson (Eds.) *The History of the Greek Language in Cyprus*, (pp. 55–66). Larnaca: Pierides Foundation.

Chang, S., Plauché, M. C., & Ohala, J. J. (2001). Markedness and consonant confusion asymmetries. In E. Hume, & K. Johnson (Eds.) *The role of speech perception in phonology*, (pp. 79–101). San Diego, CA: Academic Press Inc.

Chang, W. (2012). *R Graphics Cookbook.* Sebastopol, CA: O'Reilly.

Chantraine, P. (1980). *Dictionnaire étymologique de la langue grecque.* Paris: Klincksieck.

Chen, M. Y., & Wang, W. (1975). Sound change: Actuation and implementation. *Language*, *51*(2), 255–281.

Chennoukh, S., Carré, R., & Lindblom, B. (1997). Locus equations in the light of articulatory modeling. *The Journal of the Acoustical Society of America*, *102*(4), 2380–2389.

Cho, T., & Ladefoged, P. (1999). Variation and universals in VOT: Evidence from 18 languages. *Journal of Phonetics*, *27*(2), pp. 207–229.

Chomsky, N. (1995). *The Minimalist program*. Cambridge, MA: MIT Press.

Clackson, J. (1994). *The linguistic relationship between Armenian and Greek*. Oxford: Blackwell.

Clackson, J., & Horrocks, G. (2007). *The Blackwell history of the Latin language*. Oxford: Blackwell.

Coleman, J. (1998). Cognitive reality and the phonological lexicon: A review. *Journal of Neurolinguistics*, *11*(3), 295–320.

Coleman, J., Baghai-Ravary, L., Pybus, J., & Grau, S. (2012). Audio BNC: the audio edition of the Spoken British National Corpus.
URL `http://www.phon.ox.ac.uk/AudioBNC`

Colvin, S. (2007). *A historical Greek reader: Mycenaean to the Koiné*. Oxford: Oxford University Press.

Colvin, S. (2010). Greek dialects in the Archaic and Classical ages. In E. Bakker (Ed.) *A companion to the Ancient Greek language*, (pp. 200–212). Oxford: Wiley Blackwell.

Comparetti, D. (1914). La iscrizione arcaica di Mantinea. *Annuario della Regia Scuola Archeologica di Atene*, *1*, 1–17.

Connell, B. (1994). Structure of labial-velar stops. *Journal of Phonetics*, *22*, 441–476.

Cook, E.-D. (1981). Athapaskan linguistics: proto-Athapaskan phonology. *Annual Review of Anthropology*, (10), pp. 253–273.

Croft, W. (2000). *Explaining language change: An evolutionary approach*. New York: Longman.

Dakin, K. (1982). *La Evolución Fonológica del Protonáhuatl*. Mexico City: Universidad Nacional Autónoma de México.

Daniloff, R., & Moll, K. (1968). Coarticulation of lip rounding. *Journal of Speech, Language and Hearing Research*, *11*(4), 707.

Darden, B. J. (1970). The fronting of vowels after palatals in Slavic. In *Proceedings of the 6th Regional Meeting*, (pp. 459–70). Chicago Linguistic Society.

De Boer, B. (2000). Self-organization in vowel systems. *Journal of Phonetics*, *28*(4), 441–465.

de Vaan, M. (2008). *Etymological Dictionary of Latin*. Leiden: Brill. URL www.brillonline.com

Denzer-King, R. (2013). *The acoustics of uvulars in Tlingit*. Master's thesis, Rutgers University, New Brunswick.

Derksen, R. (2008). *Etymological dictionary of the Slavic inherited lexicon*. Leiden: Brill. URL www.brillonline.com

Desnitskaja, A. B. (1968). *Albanskij Jazyk i Ego Dialekty*. Leningrad: Nauka.

Devine, A. M., & Stephens, L. D. (1977). *Two studies in Latin phonology*. Saratoga: Anma Libri.

DiCanio, C., Nam, H., Whalen, D. H., Bunnell, H. T., Amith, J. D., & García, R. C. (2013). Using automatic alignment to analyze endangered language data: Testing the viability of untrained alignment. *The Journal of the Acoustical Society of America*, *134*(3), pp. 2235–2246.

Dorman, M. F., Studdert-Kennedy, M., & Raphael, L. J. (1977). Stop-consonant recognition: Release bursts and formant transitions as functionally equivalent, context-dependent cues. *Perception & Psychophysics*, *22*(2), pp. 109–122.

Dubois, L. (1986). Recherches sur le dialecte arcadien. *Bibliothèque des Cahiers de l'Institut de linguistique de Louvain*, (33-35).

Duhoux, Y. (2007). La lettre et quelques problèmes connexes en arcadien archaïque (IG V 2.262). *Kadmos*, *45*(1-2), 20–68.

Dunnett, R. (1970). Thessalian κίς. *Glotta*, *48*, 88.

Eckert, P. (1988). Adolescent social structure and the spread of linguistic change. *Language in Society*, *17*(2), 183–207.

Elcock, W. D. (1975). *The Romance languages*. London: Faber & Faber.

Engstrand, O., Björsten, S., Bruce, G., & Eriksson, A. (1998). Phonetic preconditions for historical sound change–evidence from the dialects. In *Proceedings of FONETIK 98, The Tenth Swedish Phonetics Conference, Stockholm University*, (pp. 24–27).

Entropic Research Laboratories, I. (1996). Entropic signal processing system (ESPS).

Ferreira, F., Henderson, J. M., Anes, M. D., Weeks, P. A., & McFarlane, D. K. (1996). Effects of lexical frequency and syntactic complexity in spoken-language comprehension: Evidence from the auditory moving-window technique. *Journal of Experimental Psychology: Learning, memory, and cognition*, *22*(2), 324.

Flanagan, J. L. (1955). A difference limen for vowel formant frequency. *The Journal of the Acoustical Society of America*, *27*(3), 613–617.

Flemming, E. S. (2002). *Auditory representations in phonology*. New York: Routledge.

Flynn, D., & Fulop, S. (2008). Dentals are grave. *Canadian Acoustics*, *36*(3), 70–71.

Fortson, B. W. (2010). *Indo-European language and culture: An introduction.* Cambridge: Cambridge University Press, 2nd ed.

Fowler, C. A. (1994). Invariants, specifiers, cues: An investigation of locus equations as information for place of articulation. *Perception & Psychophysics*, *55*(6), 597–610.

Fowler, C. A., & Housum, J. (1987). Talkers' signaling of "new" and "old" words in speech and listeners' perception and use of the distinction. *Journal of Memory and Language*, *26*(5), 489–504.

Frisk, H. (1972). *Griechisches etymologisches Wörterbuch.* Heidelberg: Winter.

Gahl, S., & Yu, A. C. (2006). Introduction to the special issue on exemplar-based models in linguistics. *The Linguistic Review*, *23*(3), 213–216.

García de León, A. (1967). La lengua de los ancianos de Jalupa, Tabasco. *Estudios de cultura Náhuatl*, *7*, pp. 267–281.

Garrett, A., & Johnson, K. (2013). Phonetic bias in sound change. In A. C. L. Yu (Ed.) *Origins of sound change: Approaches to phonologization*, (pp. 51–97). Oxford: Oxford University Press.

Gerhardt, L. (1983). *Beiträge zur Kenntnis der Sprachen des nigerianischen Plateaus.* Glückstadt: Augustin.

Goldstein, L. (1995). Lendu consonants and the role of overlapping gestures in sound change: Comments on Demolin. In B. Connell, & A. Arvaniti (Eds.) *Papers in laboratory phonology IV: Phonology and phonetic evidence*, (pp. 386–392). Cambridge: Cambridge University Press.

Grossberg, S., & Myers, C. W. (2000). The resonant dynamics of speech perception: Interword integration and duration-dependent backward effects. *Psychological Review*, *107*(4), 735.

Guarducci, M. (1937). Un giudizio del santuario di alea a mantinea. *Studi e materiali di storia delle religione*, *13*, 57–67.

Gui, M. C. (2005). *The phonology of Guangzhou Cantonese.* LINCOM studies in Asian linguistics, 56. Munich: Lincom Europa.

Guion, S. G. (1996). *Velar palatalization: Coarticulation, perception, and sound change.* Ph.D. thesis, University of Texas, Austin, Austin, TX.

Guion, S. G. (1998). The role of perception in the sound change of velar palatalization. *Phonetica*, *55*, pp, 18–52.

Haggard, M. (1973). Abbreviation of consonants in English pre- and post-vocalic clusters. *Journal of Phonetics*, *1*, pp. 9–25.

Hale, M. (2012). Foundational concepts in the scientific study of sound change. In *The Initiation of Sound Change: Perception, Production, and Social Factors*. Amsterdam: John Benjamins Publishing.

Hale, M., Kissock, M., & Reiss, C. (2007). Microvariation, variation, and the features of universal grammar. *Lingua*, *117*(4), 645–665.

Hall, D. C. (2010). On the realization of contrastive labial place. In *Proceedings of the 2010 meeting of the Canadian Linguistic Association*. Canadian Linguistic Association.

Hall-Lew, L. (2006). Glides harmonize, too: Evidence from Akan dialects. *Stanford University, MS.*.

Hardcastle, W. J., & Roach, P. J. (1979). An instrumental investigation of coarticulation in stop consonant sequences. *Current issues in the Phonetic Sciences*, *1*, pp. 531–540.

Hardison, D. M. (1999). Bimodal speech perception by native and nonnative speakers of English: Factors influencing the McGurk effect. *Language Learning*, *49*(1), 213–283.

Hargus, S. (2010). Athabaskan phonetics and phonology. *Language and Linguistics Compass*, *4*(10), 1019–1040.

Harrington, J. (2012). The coarticulatory basis of diachronic high back vowel fronting. In J. Solé, & D. Recasens (Eds.) *The initiation of sound change: Perception, production and social factors*, (pp. 103–122). Amsterdam Studies in the Theory and History of Linguistic Science.

Harrington, J., Kleber, F., & Reubold, U. (2008). Compensation for coarticulation, /u/-fronting, and sound change in standard southern British: An acoustic and perceptual study. *The Journal of the Acoustical Society of America*, *123*(5), 2825–2835.

Hasler, J. A. (1958). Dos etapas fonemáticas en el pipil de Tuztla. *Archivos Nahuas*, *1*(2), pp. 168–172.

Hawkins, S., & Warren, P. (1994). Phonetic influences on the intelligibility of conversational speech. *Journal of Phonetics*, *22*(4), 493–511.

Hawks, J. W. (1994). Difference limens for formant patterns of vowel sounds. *The Journal of the Acoustical Society of America*, *95*(2), 1074–1084.

Hay, J., & Drager, K. (2010). Stuffed toys and speech perception. *Linguistics*, *48*(4), 865–892.

Hay, J., Pierrehumbert, J., & Beckman, M. (2004). Speech perception, well-formedness, and the statistics of the lexicon. *Phonetic Interpretation: Papers in Laboratory Phonology VI, Cambridge University Press, Cambridge*, (pp. 58–74).

Heid, S., & Hawkins, S. (2000). An acoustical study of long-domain /r/ and /l/ coarticulation. In *Proceedings of the 5th seminar on speech production: Models and data*, (pp. 77–80). Citeseer.

Heiermeier, A. (1980). *Die indogermanische orientierte ursprachliche Konzeption *kuendhro-, *kuondhro-, *kuondhna-*. Dublin: Dublin University Press.

Heinrich, A., Flory, Y., & Hawkins, S. (2010). Influence of English r-resonances on intelligibility of speech in noise for native English and German listeners. *Speech Communication*, *52*(11), pp. 1038–1055.

Heinrich, A., Schneider, B. A., & Craik, F. I. M. (2008). Investigating the influence of continuous babble on auditory short-term memory performances. *The Quarterly Journal of Experimental Psychology*, *61*(5), pp. 735–751.

Heinrichs, J. (2015). Military integration in late archaic Arkadia: New evidence from a bronze pinax (ca. 500 bc) of the Lykaion. In W. Heckel, S. Müller, & G. Wrightson (Eds.) *Many Faces of War in the Ancient World*. Cambridge Scholars Publishing.

Herbert, R. K. (1975). Reanalyzing prenasalized consonants. *Studies in African Linguistics*, *6*(2), 105–123.

Heubeck, A. (1963). Tinwasijo und Verwandtes. *Minos*, (8).

Hoffner, H., & Melchert, C. (2008). *A Grammar of the Hittite Language*. Winona Lake, IN: Eisenbrauns.

Hombert, J.-M., Medjo, P., & Nguema, R. (1989). Les fangs sont-ils bantu? *Pholia*, *4*, 133.

Hooker, J. T. (1980). *Linear B: An introduction*. Bristol: Bristol Classical Press.

Horrocks, G. (2010). *Greek: A History of the Language and its Speakers*. Oxford: Wiley-Blackwell.

Houde, R. A. (1968). *A study of tongue body motion during selected speech sounds.*. Ph.D. thesis, Speech Communication Research Laboratory, University of California Santa Barbara.

Howe, D., & Fulop, S. (2005). Acoustic features in Athabascan. Unpublished manuscript, University of Calgary and California State University, Fresno.

Iskarous, K., Fowler, C. A., & Whalen, D. H. (2010). Locus equations are an acoustic expression of articulator synergy. *The Journal of the Acoustical Society of America*, *128*(4), 2021–2032.

Jacques, D. (1990). *Generative and non-linear phonology*. London: Longman.

Jakobson, R., Fant, G., & Halle, M. (1951). *Preliminaries to speech analysis: The distinctive features and their correlates*. Cambridge, MA: MIT Press.

Jakobson, R., & Halle, M. (1956). *Fundamentals of language*. The Hague, NL: Mouton de Gruyter.

Jakobson, R., & Waugh, L. R. (1979). *The sound shape of language*. Bloomington, IN: Indiana University Press.

Janson, T. (1986). Sound change in perception: An experiment. In J. J. Ohala, & J. J. Jaeger (Eds.) *Experimental Phonology*, (pp. 253–60). Orlando: Academic Press Inc.

Jasanoff, J. H. (1989). Old Irish *bé* 'woman'. *Ériu*, *40*, 135–141.

Johnson, K. (2003). *Acoustic and Auditory Phonetics*. Oxford: Blackwell, 2 ed.

Johnson, K., Sherman, V., & Sherman, S. G. (2011). *Acoustic and Auditory Phonetics*. Hoboken, NJ: Wiley-Blackwell, 3rd ed.

Johnstone, B. (1999). Uses of Southern-sounding speech by contemporary Texas women. *Journal of Sociolinguistics*, *3*(4), 505–522.

Karlgren, B. (1923). *Analytic Dictionary of Chinese and Sino-Japanese*. Courier Corporation.

Kavitskaya, D. (2002). *Compensatory lengthening: Phonetics, phonology, diachrony*. New York: Routledge.

Kawasaki-Fukumori, H. (1992). An acoustical basis for universal phonotactic constraints. *Language and Speech*, *35*(1, 2), 73–86.

Keating, P. (1993). Phonetic representation of palatalization versus fronting. *UCLA Working Papers in Phonetics*, (85).

Keating, P., & Lahiri, A. (1993). Fronted velars, palatalized velars, and palatals. *Phonetica*, *50*, 73–101.

Kewley-Port, D., & Watson, C. S. (1994). Formant-frequency discrimination for isolated English vowels. *The Journal of the Acoustical Society of America*, *95*(1), 485–496.

Kim, R. (2000). To drink, in Anatolian, Tocharian, and Proto-Indo-European. *Historische Sprachforschung*, (pp. 151–170).

Kimball, S. E. (1999). *Hittite historical phonology*. Innsbruck: Innsbrucker Beiträge zur Sprachwissenschaft.

Kingston, J. (1992). The phonetics and phonology of perceptually motivated articulatory covariation. *Language and Speech*, *35*(1, 2), 99–113.

Klatt, D. H. (1975). Voice onset time, frication, and aspiration in word-initial consonant clusters. *Journal of Speech, Language, and Hearing Research*, *18*(4), 686–706.

Kloekhorst, A. (2008). Studies in Lycian and Carian phonology and morphology. *Kadmos*, *47*, 117–146.

Kloekhorst, A. (2014). *Etymological dictionary of the Hittite inherited lexicon*. Leiden: Brill.
URL
`http://iedo.brillonline.nl/dictionaries/lemma.html?id=10942`

Krull, D. (1988). *Acoustic properties as predictors of perceptual responses: A study of Swedish voiced stops*. Ph.D. thesis.

Kuehn, D. P., & Moll, K. L. (1976). A cineradiographic study of VC and CV articulatory velocities. *Journal of Phonetics*, (4), pp. 303–320.

Kuiper, F. B. J. (1968). Pre-Hellenic labiovelars? *Lingua*, *21*, 269–277.

Kümmel, M. (2015). The role of typology in historical phonology. In P. Honeybone, & J. Salmons (Eds.) *The Oxford handbook of historical phonology*, (pp. 121–132). Oxford: Oxford University Press.

Labov, W. (1969). Contraction, deletion, and inherent variability of the English copula. *Language*, *45*(4), 715–762.

Labov, W. (1972). Some principles of linguistic methodology. *Language in Society*, *1*(01), 97–120.

Labov, W. (1982). Building on empirical foundations. vol. 24. John Benjamins Publishing.

Ladefoged, J. (1964). A phonetic study of West African languages: An auditory-instrumental survey. *The West African Languages Survey*.

Ladefoged, P. (1982). *A Course in Phonetics*. New York: Harcourt Brace Jovanovich, 2 ed.

Ladefoged, P. (1997). Linguistic phonetic descriptions. In W. Hardcastle, & J. Laver (Eds.) *The Handbook of Phonetic Sciences*, (pp. 589–618). Oxford: Blackwell.

Ladefoged, P. (2001). *Vowels and Consonants*. Oxford: Blackwell.

Ladefoged, P., Ladefoged, J., & Everett, D. (1996). Phonetic structures of Banawá, an endangered language. In *UCLA Working Papers in Phonetics: Fieldwork Studies of Targeted Languages*, vol. 93, (pp. pp. 47–66).

Ladefoged, P., & Maddieson, I. (1996). *The sounds of the world's languages*. Oxford: Blackwell.

Leer, J. (2005). How stress shapes the stem-suffix complex in Athabaskan. *Amsterdam studies in the theory and history of linguistic science 4*, *269*, 277.

Lejeune, M. (1972). *Phonétique historique du mycénien et du grec ancien*. Paris: Klincksieck.

Lillo, A. (1988). On the Reading of IG V 2, 262, 23. *Zeitschrift für Papyrologie und Epigraphik*, *73*, 86–88.

Lindblom, B. (1963a). *On vowel reduction*. Ph.D. thesis, The Royal Institute of Technology, Speech Transmission Laboratory, Stockholm.

Lindblom, B. (1963b). Spectrographic study of vowel reduction. *The Journal of the Acoustical Society of America*, *35*(11), 1773–1781.

Lindblom, B. (1986). Phonetic universals in vowel systems. *Experimental Phonology*, (pp. 13–44).

Lindblom, B. (1990). Explaining phonetic variation: A sketch of the H&H theory. In W. Hardcastle, & A. Marchal (Eds.) *Speech production and speech modelling*, (pp. 403–439). Dordrecht: Kluwer.

Lindblom, B. (1992). Phonetic content in phonology. In *Phonologica 1988: Proceedings of the 6th International Phonology Meeting*, (pp. 181–196). Cambridge: Cambridge University Press.

Lindblom, B. (1998). An articulatory perspective on the locus equation. *Behavioral and Brain Sciences*, *21*(02), 274–275.

Lindblom, B., & Sussman, H. M. (2012). Dissecting coarticulation: How locus equations happen. *Journal of Phonetics*, *40*(1), 1–19.

Lisker, L., & Abramson, A. S. (1964). A cross-language study of voicing in initial stops: Acoustical measurements. *Word*, *20*(3), pp. 384–422.

Lisker, L., & Abramson, A. S. (1967). Some effects of context on voice onset time in English stops. *Language and speech*, *10*(1), 1–28.

Löfqvist, A. (1999). Interarticulator phasing, locus equations, and degree of coarticulation. *The Journal of the Acoustical Society of America*, *106*(4), 2022–2030.

Löfqvist, A., & Gracco, V. L. (1994). Tongue body kinematics in velar stop production: Influences of consonant voicing and vowel context. *Phonetica*, *51*(1-3), pp. 52–67.

Lubotsky, A. (1989). Against a Proto-Indo-European phoneme *a. In T. Vennemann (Ed.) *The new sound of Indo-European*, (pp. 53–66). Berlin: Mouton de Gruyter.

Lynch, J. (2002). The Proto-Oceanic labiovelars: Some new observations. *Oceanic Linguistics*, *41*(2), 310–362.

Macaulay, M. A. (1996). *A grammar of Chalcatongo Mixtec*, vol. 127. Berkeley: University of California Press.

MacEacher, M. R., Kern, B., & Ladefoged, P. (1996). Wari' phonetic structure. In *UCLA Working Papers in Phonetics: Fieldwork Studies of Targeted Languages*, vol. 93, (pp. pp. 25–66).

Maddieson, I. (1996). Phonetic universals. In *UCLA Working Papers in Phonetics*, vol. 92, (pp. 160–178).

Maddieson, I., Bessell, N., & Smith, C. L. (1996). A preliminary report on the phonetics of Tlingit. In *UCLA Working Papers in Phonetics: Fieldwork Studies of Targeted Languages*, vol. 93, (pp. pp. 125–147).

Maddieson, I., & Disner, S. F. (1984). *Patterns of sounds*. Cambridge: Cambridge University Press.

Maddieson, I., Spajić, S., Sands, B., & Ladefoged, P. (1993). The phonetic structures of Dahalo. In *UCLA Working Papers in Phonetics: Fieldwork Studies of Targeted Languages*, vol. 84, (pp. pp. 25–66).

Magen, H. S. (1997). The extent of vowel-to-vowel coarticulation in English. *Journal of Phonetics*, *25*(2), 187–205.

Mann, V. A. (1986). Phonological awareness: The role of reading experience. *Cognition*, *24*(1-2), 65–92.

Mann, V. A., & Repp, B. H. (1980). Influence of vocalic context on perception of the s-ʃ distinction. *Perception & Psychophysics*, *28*(3), 213–228.

Marchand, J. W. (1973). *The sounds and phonemes of Wulfila's Gothic*, vol. 25. The Hague: Gruyter.

Martirosyan, H. K. (2008). *Etymological dictionary of the Armenian inherited lexicon*. Leiden: Brill.
URL www.brillonline.com

Matasović, R. (2009). *Etymological Dictionary of Proto-Celtic*. Leiden: Brill.
URL www.brillonline.com

Mayrhofer, M. (1986). *Indogermanische Grammatik*. Heidelberg: Winter.

McDonough, J., & Ladefoged, P. (1993a). Navajo stops. In *UCLA Working Papers in Phonetics: Fieldwork Studies of Targeted Languages*, vol. 84, (pp. 151–164).

McDonough, J., & Ladefoged, P. (1993b). The phonetic structures of Hadza. In *UCLA Working Papers in Phonetics: Fieldwork Studies of Targeted Languages*, vol. 84, (pp. 67–88).

McDorman, R. E. (1999). *Labial instability in sound change: Explanations for the loss of P*. Chicago: Organizational Knowledge Press.

McGuire, G., & Babel, M. (2012). A cross-modal account for synchronic and diachronic patterns of /f/ and /ϑ/ in english. *Laboratory Phonology*, *3*(2), 251–272.

Meier-Brügger, M. (2003). *Indo-European Linguistics*. Berlin: De Gruyter.

Meillet, A. (1894). *MLS*, vol. 8, chap. De quelques difficultés de la théorie des gutturales indo-européennes, (pp. 277–304). E. Bouillon.

Melchert, C. (2004). *The Cambridge Encyclopedia of the World's Ancient Languages*, chap. Lycian, (pp. 591–600). Cambridge: Cambridge University Press.

Melchert, C. (2012). Luvo-Lycian dorsal stops revisited. In R. Sukač, & O. Šefčík (Eds.) *The sound of Indo-European 2: papers on Indo-European phonetics, phonemics and morphophonemics*. München: Lincom Europa.

Melchert, H. C. (1987). PIE velars in Luvian. In C. Watkins (Ed.) *Studies in Memory of Warren Cowgill*, (pp. 182–204). Berlin: De Gruyter.

Melchert, H. C. (1989). New Luvo-Lycian isoglosses. *Historische Sprachforschung*, *102*, 23–45.

Melchert, H. C. (1994). *Anatolian historical phonology*. Amsterdam: Rodopi.

Mermelstein, P. (1978). Difference limens for formant frequencies of steady-state and consonant-bound vowels. *The Journal of the Acoustical Society of America*, *63*(2), 572–580.

Mielke, J. (2012). A phonetically based metric of sound similarity. *Lingua*, *122*(2), 145–163.

Miller, G. A., & Nicely, P. E. (1955). An analysis of perceptual confusions among some English consonants. *The Journal of the Acoustical Society of America*, *27*(2), pp. 338–352.

Mitterer, H. (2006). On the causes of compensation for coarticulation: Evidence for phonological mediation. *Perception & Psychophysics*, *68*(7), 1227–1240.

Monzón, C., & Seneff, A. R. (1984). Notes on the Nahuatl phonological change k^{w} > b. *International Journal of American Linguistics*, *50*(4), 456–462.

Moore, E., & Podesva, R. (2009). Style, indexicality, and the social meaning of tag questions. *Language in Society*, *38*(4), 447–485.

Mooshammer, C. (1992). *Artikulatorische Untersuchung mit EMA–Die Zungenbewegung bei der Produktion von VCV-Sequenzen mit velarer Konsonanz und langem und kurzem Erstvokal*. Master's thesis, University of Munich.

Morais, J., Alegría, J., & Content, A. (1987). The relationships between segmental analysis and alphabetic literacy: An interactive view. *Cahiers de psychologie cognitive*, *7*(5), 415–438.

Morais, J., Cary, L., Alegria, J., & Bertelson, P. (1979). Does awareness of speech as a sequence of phones arise spontaneously? *Cognition*, *7*(4), 323–331.

Morpurgo Davies, A. (2009). Greek and Indo-European semi-consonants: Mycenaean *u* and *w*. *Minos: Revista de Filología Egea*, *12*, 80–121.

Neumann, G. (2007). *Glossar des Lykischen*. No. 21 in Dresdner Beiträge zur Hethitologie. Wiesbaden: Harrassowitz.

Nord, L., & Sventelius, E. (1979). Analysis and prediction of difference limen data for formant frequencies. *STL-QPSR*, *20*(3), 60–72.

Ohala, J., & Lorentz, J. (1977). The story of [w]: An exercise in the phonetic explanation for sound patterns. In *Proceedings of the 3rd Annual Meeting of the Berkeley Linguistics Society*, vol. 3.

Ohala, J. J. (1971). The role of physiological and acoustic models in explaining the direction of sound change. *Project on Linguistic Analysis Reports*, *15*(15), 25–40.

Ohala, J. J. (1974). Phonetic Explanation in Phonology. In A. Bruck, R. Fox, & M. W. La Gally (Eds.) *Papers from the Parasession on Natural Phonology*, (pp. 251–254). Chicago: Chicago Linguistic Society.

Ohala, J. J. (1978). Southern Bantu vs. the world: the case of palatalization of labials. In *Proceedings of the Annual Meeting of the Berkeley Linguistics Society*, vol. 4, (pp. 370–386).

Ohala, J. J. (1989). Sound change is drawn from a pool of synchronic variation. In L. E. Breivik, & E. H. Jahr (Eds.) *Language change: Contributions to the study of its causes*, (pp. 173–198). Berlin: Mouton de Gruyter.

Ohala, J. J. (1990a). The phonetics and phonology of aspects of assimilation. *Papers in Laboratory Phonology*, *1*, 258–275.

Ohala, J. J. (1990b). What's cognitive, what's not in sound change. *International Symposium on 'Diachrony within synchrony', University of Duisburg*.

Ohala, J. J. (1993). Sound change as nature's speech perception experiment. *Speech Communication*, *13*(1), 155–161.

Ohala, J. J. (2005). A figure of speech: A Festschrift for john laver. In W. J. Hardcastle, & J. M. Beck (Eds.) *Phonetic explanations for sound patterns*, (pp. 23–38). London: Erlbaum.

Ohala, M., & Ohala, J. J. (1992). Phonetic universals and Hindi segment duration. In *ICSLP*, vol. 92, (pp. 831–834).

Olive, J. P., Greenwood, A., & Coleman, J. (1993). *Acoustics of American English speech*. New York: Springer Science & Business Media.

Pagliuca, W., & Mowrey, R. (1980). On certain evidence for the feature [grave]. *Language*, (pp. 503–514).

Pagliuca, W., & Mowrey, R. (1987). Articulatory evolution. In *Seventh International Conference on Historical Linguistics*, (pp. 459–472).

Palmer, L. R. (1966). Some points for discussion. In *Cambridge Colloquium on Mycenaean Studies*, (pp. 275–284).

Palmer, L. R. (1980). *The Greek Language*. London: Faber & Faber.

Panayatou, A. (2007). Arcado-Cypriot. In A.-F. Christídīs, M. Arapopoulou, & M. Chrítī (Eds.) *A history of ancient Greek: From the beginnings to late antiquity*, (pp. 417–426). Cambridge: Cambridge University Press.

Pedersen, H. (1945). *Lykisch und Hethitisch*. Copenhagen: I kommission hos ejnar Munksgaard.

Perkell, J. S., & Cohen, M. H. (1989). An indirect test of the quantal nature of speech in the production of the vowels /i/, /a/ and /u/. *Journal of Phonetics*, *17*(1-2), 123–133.

Perkell, J. S., & Matthies, M. L. (1992). Temporal measures of anticipatory labial coarticulation for the vowel /u/: Within- and cross-subject variability. *The Journal of the Acoustical Society of America*, *91*(5), 2911–2925.

Picard, M. (1984). On the Naturalness of Algonquian ɬ. *International Journal of American Linguistics*, *50*(4), 424–437.

Pierrehumbert, J. (2001). Exemplar dynamics: Word frequency, lenition and contrast. In J. L. Bybee, & P. J. Hopper (Eds.) *Frequency and the Emergence of Linguistic Structure*, (pp. 137–157). Amsterdam: John Benjamins Publishing.

Plauché, M., Delogu, C., & Ohala, J. J. (1997). Asymmetries in consonant confusion. *Fifth European Conference on Speech Communication and Technology, EUROSPEECH*.

Pokorny, J. (1959). *Indogermanisches Etymologisches Wörterbuch*. Tübingen: Francke.

Ponelis, F. (1974). On the dynamics of velarization and labialization: Some Bantu evidence. *Studies in African Linguistics*, *5*(1).

Port, R. F. (2010). Rich memory and distributed phonology. *Language Sciences*, *32*(1), 43–55.

Probert, P. (2008). Mycenaean o- is accusative; jo- is nominative. *Glotta*, *84*(1-4), 126–168.

Probert, P. (2010). Phonology. In E. Bakker (Ed.) *A Companion to the Ancient Greek Language*, (pp. 83–103). Wiley Blackwell.

Pullum, G. K., & Ladusaw, W. A. (1996). *Phonetic symbol guide*. Chicago: University of Chicago Press.

Read, C., Zhang, Y.-F., Nie, H.-Y., & Ding, B.-Q. (1986). The ability to manipulate speech sounds depends on knowing alphabetic writing. *Cognition*, *24*(1), 31–44.

Recasens, D. (1985). Coarticulatory patterns and degrees of coarticulatory resistance in Catalan CV sequences. *Language and Speech*, *28*(2), 97–114.

Recasens, D. (1999). Lingual coarticulation. In D. Recasens, W. Hardcastle, & N. Hewlett (Eds.) *Coarticulation: Theory, data and techniques*, vol. 4, (pp. 80–100). Cambridge: Cambridge University Press.

Recasens, D. (2012). A phonetic interpretation of the sound changes affecting dark /l/ in Romance. In M.-J. Solé, & D. Recasens (Eds.) *The Initiation of Sound Change: Perception, production, and social factors*. Amsterdam: John Benjamins Publishing.

Rice, K. (1989). *A grammar of Slave*. No. 5 in Mouton Grammar Library. New York: Mouton de Gruyter.

Risch (1979). Die griechischen Dialekte im 2. vorchristlichen Jarhtausend. *SMEA*, *20*, 91–104.

Rix, H. (2001). *Lexikon der indogermanischen Verben*. Wiesbaden: Reichert.

Rosner, B. S., López-Bascuas, L. E., García-Albea, J. E., & Fahey, R. P. (2000). Voice-onset times for Castilian Spanish initial stops. *Journal of Phonetics*, *28*(2), 217–224.

Ross, M. (2011). Proto-oceanic $*k^w$. *Oceanic Linguistics*, *50*(1), 25–49.

Schrijver, P. (1991). *The reflexes of the Proto-Indo-European laryngeals in Latin*. Leiden Studies in Indo-European. Amsterdam: Rodopi, 2 ed.

Schwyzer, E. (1939). *Griechische Grammatik*.

Sen, R. (2012). Reconstructing phonological change: Duration and syllable structure in Latin vowel reduction. *Phonology*, *29*(03), 465–504.

Shaul, D. (2010). Languages of the opatería. In D. A. Yetman (Ed.) *The Ópatas: in search of a Sonoran people*, (pp. pp. 259–272). University of Arizona Press.

Sheets, G. A. (1975). Palatalization in Greek. *Indogermanische Forschungen*, *80*, 118.

Shepard, R. N. (1972). Psychological representation of speech sounds. In E. E. David, & P. B. Denes (Eds.) *Human communication: A unified view*, (pp. 67–113). New York: McGraw-Hill.

Shockey, L. (1977). Perceptual test of a phonological rule. *Haskins Laboratories Quarterly Progress Report*, (pp. 147–50).

Sicoli, M., & Kaufman, T. (2010). Zapotec and Chatino Survey Archive [online]. *Max Planck Institute, Archive for Linguistic Resources*. (Access restricted and available on request).
URL
`http://corpus1.mpi.nl/ds/imdi_browser?openpath=MPI663149%23`

Sihler, A. L. (1995). *New comparative grammar of Greek and Latin*. Oxford: Oxford University Press.

Silverman, D. (2006). The diachrony of labiality in Trique, and the functional relevance of gradience and variation. *Laboratory phonology 8: Varieties of phonological competence*, (pp. 133–152).

Sjölander, K., & Beskow, J. (2000). Wavesurfer - an open source speech tool. In *INTERSPEECH*, (pp. 464–467).

Smeets, R. (1986). On Ubykh Circassian. *Studia Caucasologica I: Proceedings of the Third Caucasian Colloquium*.

Solé, M. J. (2014). The perception of voice-initiating gestures. In J. Harrington, M. Stevens, & J. Cole (Eds.) *The Relationship Between the Origin and Spread of Sound Change*, vol. 5 of *Laboratory Phonology*, (pp. 37–68). Mouton de Gruyter.

Steele, S. (1979). Uto-Aztecan: An assessment for historical and comparative linguistics. In L. Campbell, & M. Mithun (Eds.) *The Languages of Native America*, (pp. 444–544). Austin, TX: University of Texas Press.

Stephens, L. D., & Woodard, R. D. (1986). The palatalization of the labiovelars in Greek: A reassessment in typological perspective. *Indogermanische Forschungen*, *91*, 129.

Stevens, K. N. (1998). *Acoustic Phonetics*. No. 30 in Current Studies in Linguistcs. Cambridge, MA: The MIT Press.

Stevens, K. N., & House, A. S. (1963). Perturbation of vowel articulations by consonantal context: An acoustical study. *Journal of Speech & Hearing Research*, *6*(2), 111–128.

Stuart-Smith, J. (2004). *Phonetics and philology: Sound change in Italic*. Oxford: Oxford University Press.

Stuart-Smith, J., Timmins, C., & Tweedie, F. (2007). Talkin' Jockney? Variation and change in Glaswegian accent. *Journal of Sociolinguistics*, *11*(2), 221–260.

Stubbs, B. D. (1995). The labial labyrinth in Uto-Aztecan. *International Journal of American Linguistics*, *61*(4), 396–422.

Stubbs, B. D. (2011). *Uto-Aztecan: A comparative vocabulary*. Shumway Family History Services.

Sturtevant, E. H. (1940). *The Pronunciation of Greek and Latin*. New Haven, CT: Linguistic Society of America, 2 ed.

Suh, Y. (2007). Consonant-labiovelar glide combinations in Spanish and Korean. In *Proceedings of ICPhS*, vol. 16, (pp. 1773–1776).

Suh, Y. (2008). The effect of secondary labialization on stop burst spectra. *The Journal of the Acoustical Society of America*, *124*(4), pp. 2556–2556.

Sussman, H. (2015). A functional role for coarticulation: A locus equation perspective. In A. Agwuele, & A. Lotto (Eds.) *Essays in Speech Processes: Language production and perception*. Sheffield: Equinox.

Sussman, H. M., McCaffrey, H. A., & Matthews, S. A. (1991). An investigation of locus equations as a source of relational invariance for stop place categorization. *The Journal of the Acoustical Society of America*, *90*(3), 1309–1325.

Sussman, H. M., & Shore, J. (1996). Locus equations as phonetic descriptors of consonantal place of articulation. *Perception & Psychophysics*, *58*(6), 936–946.

Suzuki, K. (1998). *A typological investigation of dissimilation*. Ph.D. thesis, University of Arizona.

Szemerényi, O. (1966). The labiovelars in Mycenaean and historical Greek. *Studi micenei ed egeo-anatolici*, (pp. 1–29).

Tabain, M. (2000). Coarticulation in CV syllables: A comparison of locus equation and EPG data. *Journal of Phonetics*, *28*(2), 137–159.

Tabain, M. (2002). Voiceless consonants and locus equations: A comparison with electropalatographic data on coarticulation. *Phonetica*, *59*(1), 20–37.

Taylor, A. R. (1963). Comparative Caddoan. *International Journal of American Linguistics*, *29*(2), pp. 113–131.

Thomason, S. G. (1986). On changes from palatalized labials to apical affricates. *International Journal of American Linguistics*, *52*(2), 182–186.

Thompson, R. (2006). Long mid vowels in Attic-Ionic and Cretan. *The Cambridge Classical Journal*, *52*, 81–101.

Thompson, R. (2010). Mycenaean Greek. In E. Bakker (Ed.) *A companion to the Ancient Greek language*, (pp. 189–199). Oxford: Wiley-Blackwell.

Traunmüller, H., & Öhrström, N. (2007). Audiovisual perception of openness and lip rounding in front vowels. *Journal of Phonetics*, *35*(2), 244–258.

Vago, R. M. (1976). More evidence for the feature [grave]. *Linguistic Inquiry*, *7*(4), 671–674.

Välimaa-Blum, R. (2009). The phoneme in cognitive phonology: Episodic memories of both meaningful and meaningless units? *CogniTextes*, (Volume 2). Revue de l'Association française de linguistique cognitive.

Vendryes, J. (1959). *Lexique étymologique de l'irlandais ancien*, vol. 1. Dublin: Dublin Institute for Advanced Studies.

Ventris, M., & Chadwick, J. (1973). *Documents in Mycenaean Greek*. Cambridge: Cambridge University Press, 2nd ed.

Vis, J. (2011). Synchrony in diachronic analysis: The interpretation of *w* in Mycenaean Greek. In E. Kitis, N. Lavidas, N. Topintzi, & T. Tsangalidis (Eds.) *Selected papers from the 19th International Symposium on Theoretical and Applied Linguistics*, (pp. 481–490). Monochromia Publishing.

Wachter, R. (1987). *Altlateinische Inschriften*. Bern: Peter Lang.

Weiss, M. (1994). Life everlasting: Latin *iūgis* 'everflowing', Greek ὑγιής 'healthy', Gothic *ajukdūþs* 'eternity' and Avestan *yauuaē̆ǰī-* 'living forever'. *Münchener Studien zur Sprachwissenschaft*, *55*, pp. 131–156.

Welmers, W. E. (1973). *African language structures*. Berkeley, CA: University of California Press.

West, P. (1999). Perception of distributed coarticulatory properties of English /l/ and /r/. *Journal of Phonetics*, *27*(4), 405–426.

Westermann, D., & Ward, I. C. (1933). *Practical phonetics for students of African languages*. London: Oxford University Press.

Whatmough, J. (1937). The development of the Indo-European labiovelars with special reference to the dialects of Italy. *Acta Jutlandica*, *9*, 45–56.

Wilson, C. (2006). Learning phonology with substantive bias: An experimental and computational study of velar palatalization. *Cognitive Science*, *30*(5), 945–982.

Wilson, W. A. A., & Bendor-Samuel, J. T. (1969). The phonology of the nominal in Dagbani. *Linguistics*, *7*(52), 56–82.

Winitz, H., Scheib, M. E., & Reeds, J. A. (1972). Identification of stops and vowels for the burst portion of /p, t, k/ isolated from conversational speech. *The Journal of the Acoustical Society of America*, *51*, 1309–1317.

Woodard, R. D. (2010). Phoinikeia grammata: An alphabet for the Greek language. In E. Bakker (Ed.) *A Companion to the Ancient Greek Language*, (pp. 25–46). Oxford: Wiley-Blackwell.

Zima, P. (1974). Remarques sur le conconsonant d'un parler Haoussa du Niger (les problémes de 'palatalisation'). *Phonetica Pragensia*, *4*, 123–132.

Zima, P. (1985). Labiovelar stops in the Djougou Dendi dialect of Songhay. *Phonetica Pragensia*, *7*, 97–103.

Zue, V. W. (1976). *Acoustic characteristics of stop consonants: A controlled study*. Ph.D. thesis, Massachusetts Institute of Technology.