

Asymmetric processing of consonant duration in Swiss German

LARA EHRENHOFER, ADAM C. ROBERTS, SANDRA KOTZOR, ALLISON WETTERLIN, AND ADITI LAHIRI

9.1 Introduction

Understanding spoken language relies on extracting information from a highly variable acoustic signal and mapping that information onto abstract representations of phonological categories. Most investigations of this process have concentrated on feature-based phoneme contrasts (e.g. the spectral information that cues place of articulation, or the durational information that cues voice and voicelessness in stop series) and their representation in the mental lexicon. Radical underspecification models of the mental lexicon, such as the Featurally Underspecified Lexicon (FUL) (Lahiri and Reetz 2002, 2010), provide a maximally efficient way to represent and recognize segments by storing only unpredictable (marked) feature information. But for phonemic distinctions like the geminate/singleton contrast, the problem of recognition shifts: the segment's phonetic duration is mapped onto a syllable structure¹ as well as onto a phoneme. To what extent is this structural distinction represented in the mental lexicon? If it is, is that representation underspecified, as feature representations appear to be? Geminates are therefore a keystone for experimental work at the interface of phonetics and phonology, as the single phonetic cue of duration is implicated in both melodic and structural levels of representation within the mental lexicon.

We investigate geminate representation through the lens of how durational information is exploited in lexical access. Prior work on Bengali (Lahiri and Marslen-Wilson 1992; Roberts et al. 2014; Kotzor et al. 2015; Kotzor et al., Chapter 8 in this volume)

¹ This is most obviously true of medial geminates, which are the main focus of our discussion, but arguably also applies to initial and final geminates when they are embedded within an utterance, as they resyllabify with syllables from adjacent words (Kraehenmann 2003).

showed that consonant duration is processed asymmetrically, suggesting that the representation of the geminate/singleton contrast is indeed structurally under-specified. However, Bengali only allows geminates in medial position, so speakers may be unable to dissociate phonetic duration from a heterosyllabic structure for phonotactic reasons, possibly influencing these asymmetric results. Swiss German differs crucially from Bengali in allowing geminates word-initially and finally as well as medially. Its speakers therefore have the ability to evaluate a consonant's phonetic duration and represent it phonologically in all positions independently of heterosyllabicity, ruling out the potential role of phonotactics in processing. We detail the phonetic and phonological characteristics of geminates cross-linguistically, and more specifically in Swiss German, in section 9.2. Section 9.3 discusses prior work on geminate representation in the mental lexicon, and lays out our research questions.

Our experiment, described in section 9.4, follows the logic of Roberts et al. (2014) and Kotzor et al. (Chapter 8). It contrasts phoneme recognition at the melodic and at the timing level by presenting listeners with auditory primes whose medial consonant duration has been altered, such that words with medial singletons are presented with geminate duration and vice versa. We investigate this manipulation's impact on lexical access through the N400, an electrophysiological component which indexes lexical access (Lau et al. 2008), and find a processing asymmetry. In section 9.5, we lay out two alternative explanations for our results: a representational account based on a structural specification for geminates in the mental lexicon, and a processing account, which suggests that the asymmetry we observe is due to a parsing mechanism (the 'CV template') which initially assigns consonants in the acoustic signal to syllable onsets.

9.2 Why geminates are special

Duration is a gradable property of the speech signal and gives rise to clear categorical perception effects in phoneme identification. For instance, the phonetic duration of voice onset time (VOT) distinguishes between phonologically voiced and voiceless stops in English (Lisker et al. 1977; Reetz and Jongman 2009). This is no different for the specific case of geminate duration: presenting participants with stimuli with graded closure durations in languages as diverse as Japanese (Amano and Hirata 2008), Turkish and Bengali (Hankamer et al. 1989), and Arabic (Obrecht 1965) shows that cross-linguistically, consonant duration is the primary cue to geminate or singleton identity. What sets geminate duration apart from duration purely at the melodic level (as in English VOT) is the fact that it simultaneously cues a structural distinction in syllabification. Processing phonetic duration in a length-contrastive system therefore requires the listener to mediate between melodic and structural levels of representation. Furthermore, the precise nature of these representations is

under considerable debate (see section 9.3). In this section, we review phonetic studies of geminates in various languages, then go on to outline some important phonetic and phonological features of geminates in Swiss German and in other languages. Since our primary interest is in how duration in the acoustic signal is mapped onto underlyingly long and short consonant categories (rather than how listeners deal with durational variation arising from morphological or prosodic factors), our discussion here focuses on lexical geminates.

9.2.1 Phonetic characteristics of geminates

In both perception and production studies, a great deal of research has focused on attempts to find secondary correlates (beyond duration) to the geminate/singleton distinction. Robust secondary correlates would tip the scales in favour of a geminate representation focused on melodic features, but such evidence has (with the possible exception of Pattani Malay, see below) been relatively unstable. In one perception study (Lahiri and Hankamer 1988), Turkish displayed a statistically significant difference in VOT across geminates and singletons. Nevertheless, cross-spliced stimuli (in which geminate stops were presented with a singleton VOT and singletons with geminate VOT) were still reliably perceived as containing geminate consonants, and a follow-up study with gated stimuli (Hankamer et al. 1989) confirmed that geminate recognition was primarily dependent on consonant duration rather than secondary cues like VOT.

Existing production studies, on the other hand, suggest that languages differ according to whether or not they have secondary acoustic characteristics that correlate robustly with the geminate/singleton distinction. Ridouane (2007) revealed a number of potential candidates (e.g. variation in fundamental frequency, variation in release duration) for secondary correlates to the distinction in word-initial, medial, and final geminates in Tashlhiyt Berber, but none of these were robustly present in all speakers and all consonant types, suggesting that these acoustic differences are a result of phonology-external factors (e.g. aerodynamics, Ohala 1983). However, in Malayalam, F_1 values in vowels preceding singleton sonorants are systematically lower than in vowels preceding geminates (Local and Simpson 1999). Extending to articulation, Payne (2005, 2006) finds statistically robust acoustic and electropalatographic (EPG) evidence of a stronger palatal articulation for geminate than singleton sonorants in Italian lexical geminates. In Malayalam and Italian, it may be the case that listeners are able to exploit robustly available supplementary acoustic cues to support the identification of geminates vs. singletons. However, as neither of these studies included a perceptual task, it is impossible to tell to which extent duration can be overridden by other cues to the geminate/singleton contrast. In some languages, geminate duration may be accompanied by secondary acoustic characteristics, but these do not seem to be robust cues that listeners can reliably exploit to make the geminate/singleton distinction. This is also true for Swiss German.

9.2.2 *Geminates in Swiss German*

Swiss German belongs to the High Alemannic branch of German and is spoken by about 4.64 million people in Switzerland (2000 census; Lewis et al. 2013), where it exists in a diglossic situation alongside Standard German, which is used for written communication and in formal situations (Ferguson 1959). Mutual intelligibility between Standard and Swiss German is low (Lewis et al. 2013), and numerous differences in their phonological systems (Kraehenmann 2003) and syntax (Schieber 1985; Glaser and Frey 2011) suggest that Swiss German is a language of its own, rather than a dialect of Standard German. There are some differences between the phonetics and phonological systems of different regional accents (e.g. discrepancies in vowel length between Grison and Thurgovian Swiss German, Kraehenmann 2003), leading scholars to refer more specifically to e.g. 'Bernese Swiss German' (Ham 2001) or 'Thurgovian Swiss German' (Kraehenmann 2003). However, the gemination facts we address in the present work hold across all dialects of Swiss German.

Many scholars (e.g. Brunner 1954; Enstrom and Spörri-Bütler 1981; Fulop 1994) have attempted to explain the Swiss German singleton/geminate contrast as a difference in fortis/lenis (tense/lax) articulation. This stems from a long tradition of attempting to bring the phonological system of Swiss German into line with an account of High German dialects (including modern-day Standard German) in which a phonological voicing distinction in stop series is realized phonetically as aspiration ([± spread glottis], Iverson and Salmons 1995). This underlying distinction in High German has been analysed by many authors (perhaps most notably by Kohler 1984) as a realization of [± fortis]; but for Swiss German, this account of consonantal length contrasts breaks down under inquiry from phonetic, phonological, and historical points of view (see Ehrenhofer 2013 for a detailed discussion).

As numerous phonetic studies of Swiss German consonants have confirmed, aspiration does not differ significantly between the purported fortis and lenis consonants of Swiss German (Enstrom and Spörri-Bütler 1981; Fulop 1994; Willi 1996); if there is voicing, it is not consistent across speakers or consonants (Brunner 1954). Figures 9.1 and 9.2 show sample waveforms of the minimal pair [vɔpə] 'honeycomb' ~ [vɔp:ə] 'coat of arms.'

In-depth phonetic studies of Thurgovian Swiss German (Kraehenmann 2001, 2003) show significant differences in duration between underlyingly long vs. short consonants, although the duration ratios differ depending on manner of articulation and position within the word: the difference in consonant duration is most pronounced word-medially, and across all manners of articulation, stops have the greatest ratio of singleton to geminate consonant duration (see Table 9.1 for a summary of the phonetic data in Kraehenmann 2001, 2003).

Initial stops in Swiss German provide further evidence that duration is the primary cue to the geminate/singleton distinction in this language. The geminate/singleton

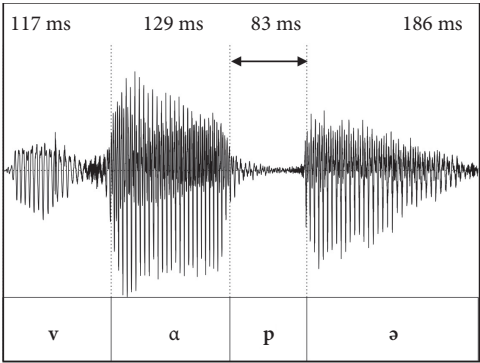


FIGURE 9.1 Sample waveform of [vʌpə] ‘honeycomb’.

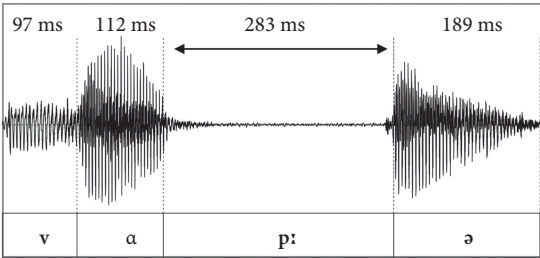


FIGURE 9.2 Sample waveform of [vʌp:ə] ‘coat of arms’.

TABLE 9.1. Mean geminate and singleton consonant durations in Swiss German in milliseconds

	Initial	Ratio	Medial	Ratio	Final	Ratio	Overall	Ratio
Singleton stop	67.8	1:1.26	64.0	1:2.4	70.1	1:1.65	67.3	1:1.77
Geminate stop	86.3		155.2		116.1		119.2	
Singleton fricative			94.0	1:1.79	122.1	1:1.3	104.0	1:1.3
Geminate fricative			169.0		158.2		135.8	
Singleton sonorant			63.4	1:1.76	102.3	1:1.23	61.5	1:1.85
Geminate sonorant			111.4		126.3		113.5	

distinction in stops is realized phonetically as a difference in the length of the silence before the burst, measured from the end of the preceding vowel or consonant segment. In some languages, secondary cues to consonant identity are available: speakers of Pattani Malay can exploit formant transitions into, and amplitude of,

TABLE 9.2. The consonant inventory of Swiss German

	Bilabial	Labiodental	Coronal	Palato-alveolar	Velar	Glottal
Nasal	m: m		n: n		ŋ:	
Plosive	p: p		t: t		k: k	
Fricative		f: f	s: s	ʃ: ʃ	x: x	h
Affricate	pf		ts	tʃ	kx	
Trill			(r)		(R)	
Lateral			l: l			

Note: The realization of /r/ differs between dialects and speakers, and is even subject to free variation (see Fleischer and Schmid's (2006) narrow transcription of Swiss German).

the following vowel (Abramson 1986b, 1987, 1991) to make this distinction, while speakers of Kelantan Malay use VOT as a cue to a consonant's geminate or singleton status (Hamzah et al. 2011). Speakers of Swiss German maintain the phonological distinction between utterance-initial voiceless singletons and geminates in production (Kraehenmann and Lahiri 2008), but are unable to make this distinction in perception (Kraehenmann 2001, 2003), suggesting that no reliable secondary cues to consonant identity are available to speakers.

Swiss German's phonological length contrast, realized primarily as a difference in duration, pervades the entire synchronic consonant system² (see Table 9.2, based on Kraehenmann 2003). There are just two serious restrictions on geminate distribution: the only geminates found word-initially are stops (e.g. [tʌŋk] 'thanks' vs. [t:ʌŋk] 'tank'), and geminate sonorants can only appear following a short vowel ([gɛ:lə] 'loudmouth', [gɛ:lə] 'yellow.masc.acc', but a word of the form *[gɛ:l:ə] is phonotactically illegal). Otherwise, consonant length is distinctive following every nucleus type: long or short vowels, e.g. [vʌp:ə] 'coat of arms', [vʌpə] 'honeycomb', [hu:p:ə] 'to honk', [hu:pə] 'cap'; branching or non-branching nucleus, e.g. [ʌlt:] 'old'. Phonetically long consonants also arise as a result of morphological concatenation at word boundaries ('fake' geminates, Kraehenmann 2003), but these lie beyond the scope of this chapter. Swiss German, unlike Italian (Payne 2005) and Cypriot Greek (Payne and Eftychiou 2006), does not have geminates resulting from prosodic factors.

9.3 Geminates in the mental lexicon

General issues with the representation of geminates are set out in detail in Kotzor et al. (Chapter 8). We reiterate here that geminate duration cannot be considered a

² There are a few exceptions: affricates and /ŋ/, which are considered to be underlyingly bisegmental, only appear as phonetically long; and /r/, which only appears as phonetically short (Kraehenmann 2003).

melodic feature. If it were, long and short consonants would be represented by two separate phonemes [C] vs. [C:] differentiated by the feature [\pm long], analogous to the distinction of oral from nasal vowels through a feature [\pm nasal] (cf. e.g. Chomsky and Halle 1968). However, the geminate/singleton distinction differs fundamentally in that it represents in the lexicon the amount of phonetic duration, rather than its presence or absence. Featural phonological distinctions may be phonetically realized through differences in duration—VOT in English is one such example—but crucially, this graded phonetic cue maps onto an underlying distinction that is represented as a presence or absence of voicing ([\pm voice]). Furthermore, consonants which are distinguished by the presence or absence of a feature tend to be affected differently (Schein and Steriade 1986), whereas geminates and singletons are always affected identically, suggesting that they share all melodic information and differ on some other dimension.

Instead, the geminate/singleton must be encoded as a structural distinction. There is controversy over whether geminates are represented using a binary (one vs. two units of representation) or privative system (zero vs. one unit), and what form this representation takes (e.g. moras vs. some other structural specification). Here we primarily address the question of a binary (specified) vs. privative (underspecified) representation, as experimental methods do not currently allow us to distinguish between moraic and other theoretical accounts of geminate representation.

To our knowledge, consonant duration is not explicitly accounted for in any model of the mental lexicon. Given experimental evidence favouring radically underspecified phonological representation in the mental lexicon, we focus here on three central questions: what underspecified representations look like, how existing experimental work supports an underspecified representation for the geminate/singleton in Bengali, and why this may or may not apply to geminates in Swiss German. Our aim is to leverage the examples of Bengali and Swiss German to foster debate on the precise nature of the specification, how this could impact models of the mental lexicon, and the division of labour between phonological representation and processing in speech recognition.

Early and influential models of lexical access, such as TRACE (McClelland and Elman 1986), assumed featurally fully specified representations. In this model, recovery from a distorted signal (e.g. through mispronunciation or phonological assimilation) is mediated by top-down influence from currently active word candidates, as well as activation or inhibition from phoneme to feature units in adjacent segments. As Gaskell and Marslen-Wilson (1996) point out, this predicts that all deviations from a target feature should be equally easy to recover from, but beginning with Lahiri and Marslen-Wilson (1991), experimental results have shown that phoneme perception often proceeds asymmetrically. To account for this type of processing asymmetry, Lahiri and Reetz (2002) elaborated the Featurally Underspecified Lexicon (FUL), which assumes privative features with radical underspecification. If the acoustic

input matches a specified feature in the mental lexicon (MATCH), processing is facilitated. If the acoustic input cannot be mapped onto a specified feature (e.g. because the information is missing or contradicts the specification; MISMATCH), processing is inhibited. If a feature is underspecified, any acoustic input can map onto it (NO-MISMATCH), again resulting in facilitated processing. The predictions of this model have been borne out repeatedly. Friedrich et al. (2006) and Friedrich et al. (2008) conducted a cross-modal semantic priming study in German, which is underspecified for [CORONAL], using audio primes whose medial, final, or initial consonant differed in place of articulation from the lexical representation. Behavioural and neurophysiological evidence showed that participants were able to map a labial acoustic signal onto an underlyingly coronal segment, but not the other way round (**Horbe* primes targets that are semantically related to *Horde* 'horde', i.e. there is a no-mismatch response, but **Prode* does not prime targets that are semantically related to *Probe* 'sample', i.e. a mismatch response is provoked). Lahiri and Reetz (2010) applied essentially the same methodology to German vowels, but derived the variation in vowel quality through suffixes which triggered umlaut in the vowel, rather than creating non-words, and again reported a processing asymmetry best explained by FUL.

Taken together, these studies suggest that processing acoustic input relies on extracting melodic information and mapping it onto an underspecified phonological representation in the mental lexicon. But to what extent does underspecified representation carry over to the processing of acoustic information that is represented not in terms of melodic but in terms of structural information? Experimental work on geminate duration, which is not a feature (Lahiri and Marslen-Wilson 1992; Kraehenmann 2003), helps answer this question.

Two major studies provide support for an underspecified representation of geminates in Bengali. In a forced-choice identification task, Lahiri and Marslen-Wilson (1992) presented participants with gated fragments of a geminate/singleton minimal pair (e.g. *kana* 'blind' vs. *kanna* 'tears'). Stimuli cut off at the third gate contained the full duration of the medial sonorant, yet even given the full duration of the medial geminate or singleton sonorant, listeners overwhelmingly classified stimuli as corresponding to a word with a medial singleton. Recognition of duration as belonging to a geminate only rose above chance level at the following gate, which also included the consonant release and the first two glottal pulses of the vowel. This asymmetric processing profile for Bengali geminate duration suggests that the geminate/singleton contrast is underspecified at the structural level, since geminate recognition depends on the availability of a cue to syllable structure (in the form of information about an upcoming vowel). Roberts et al. (2014) replicated this processing asymmetry for Bengali geminates by demonstrating that speakers presented with non-word stimuli with a phonetically long medial consonant were able to access lexical forms with underlying singletons, but were not able to map a non-word with a phonetically short medial consonant onto an underlying geminate. On the basis of this neurolinguistic

and behavioural evidence for asymmetric duration processing, Kotzor et al. (Chapter 8) argue that duration in Bengali is privative (\emptyset vs. μ): singletons are unspecified for length, but geminates are specified for mora possession.

Our further experimental investigations in Swiss German add to the debate on geminate representation in two ways. Firstly, Swiss German geminates have been analysed as non-moraic because they are present initially, yet do not contribute to syllable weight (Kraehenmann 2003). If we nonetheless observe a processing asymmetry, this spurs discussion on alternative non-moraic structural units that geminates might be specified for. Secondly, Bengali phonotactics only licenses a phonological contrast in duration in word-medial position, possibly contributing to the observed processing asymmetry. Recall that stimuli in the crucial conditions of Lahiri and Marslen-Wilson's (1992) task were truncated at the end of the medial consonant, forming CVC sequences with variation in the duration of the final consonant. Since Bengali does not allow syllables with a final geminate, speakers may not have the representational capacity to encode durational variation in coda consonants. If so, there is a chance that the finding that Bengali speakers are strongly biased towards perceiving consonants as singletons (and therefore suggesting an underspecified representation of geminates) may be an artefact of the phonotactic limitations of Bengali. Since Swiss German allows geminates in all positions within the word, it provides a unique test of questions about the processing and representation of geminate duration. If the processing asymmetry observed in Bengali geminates is an artefact of phonotactics, rather than representation, Swiss German speakers should evaluate durational information symmetrically. If, however, Swiss Germans display the same asymmetry in processing as Bengali speakers, this would suggest that geminate representation is asymmetric cross-linguistically, ruling out the possibility of phonotactics as a reason for the Bengali asymmetry, and that the geminate/singleton contrast may be underspecified at a structural level.

9.4 Asymmetric processing of consonant duration in Swiss German

Our experiment examines how speakers of Swiss German extract duration information from an acoustic signal and integrate this information in lexical access, and explores the implications of our findings for models of the mental lexicon.

The design follows techniques pioneered in, amongst many others, Lahiri and Reetz (2002, 2010); Friedrich, Eulitz, and Lahiri (2006); Friedrich, Lahiri, and Eulitz (2008); and Roberts et al. (2014). We exploit the robust and widespread finding that processing is facilitated for words that are presented simultaneously with or after (Meyer and Schvaneveldt 1971; Bentin et al. 1985) semantically related words, i.e. the prime *dog* facilitates the processing of the target *wolf*, but the prime *nail* does not. Facilitated processing is reflected behaviourally in shorter reaction times (Neely 1976) and neurophysiologically in a decrease in the amplitude of a negative-going electrophysiological

component known as the N400. First observed in the context of sentence processing (in semantically anomalous sentences like 'He likes his coffee with cream and *socks*', Kutas and Hillyard 1980, 1984), the N400 is now understood to index lexical access, semantic integration, and prediction or expectedness (Kutas and Federmeier 2000; Lau et al. 2008). The advantage of using electrophysiological measures like event-related potentials (ERPs) alongside behavioural measures is that it provides precise, millisecond-by-millisecond insights into processing as it unfolds, whereas a response time indicates the overall outcome of the decision subjects make as a result of that processing.

We manipulated the consonant duration of aurally presented primes, and measured reaction times and ERPs to semantically related and unrelated targets for an indication of the success of lexical access. In the crucial conditions of our experiment, Swiss German speakers heard non-word primes based on words whose medial geminate or singleton had been manipulated to display the opposite duration: words containing medial singletons were presented with medial geminate duration (e.g. [ʃtʊpə] 'living room' was presented as *[ʃtʊp:ə], which is not a word in Swiss German), and vice versa (e.g. [sʊp:ə] 'soup' was presented as the non-word *[sʊpə]). Participants then responded to semantically related visual targets (for these examples, *Kammer* 'chamber' and *Bouillon* 'broth', respectively). Where recognition is facilitated, this indicates that the acoustic input has been successfully mapped onto the underlying representation. Where recognition is inhibited, this suggests that the manipulation of medial consonant duration led to a failure in this mapping.

We see three possible outcomes to our experiment. Priming may be equally reduced in responses to targets paired with both singleton-to-geminate and geminate-to-singleton non-words (i.e. lexical access to the prime is hindered, resulting in a reduced facilitation of access to the target). This would suggest a symmetric representation, i.e. that Swiss German listeners immediately evaluate durational information and map it to a binary opposition that matches the two phonetic durations (so a phonetically short duration could only be mapped onto a singleton representation and a phonetically long duration onto a geminate representation). If phonetic duration does not match the underlying specification, lexical access (reflected in priming) would fail. This result would lend support to the hypothesis that the asymmetric geminate processing found in gated Bengali stimuli (Lahiri and Marslen-Wilson 1992) is an artefact of phonotactics rather than representation. Alternatively, we may find priming across the board (regardless of manipulation). Studies using primes with melodic deviations from real words (e.g. **domato* or **somato* for *tomato*, Bölte and Coenen 2000) have found a cline in the tolerance for deviation that the degree of priming reflects, and it is possible (though in our opinion unlikely) that durational variation in Swiss German might be tolerated better than in Bengali. The third possible outcome is that Swiss German will show the same processing asymmetry as Bengali, with a reduced priming effect in response

to geminate-to-singleton non-words, but not singleton-to-geminate non-words (i.e. show a privative contrast). This would indicate that although Swiss German encodes durational information in all positions within the word, speakers are biased towards interpreting consonants as singletons, suggesting that the representation of the geminate/singleton contrast in Swiss German, as in Bengali, is underspecified at the structural level.

9.4.1 Experiment design and materials

Our experimental design follows that of Roberts et al. (2014) (also discussed in Chapter 8). Audio primes and visual targets were paired in a blocked design (real word vs. non-word; based on a singleton or based on a geminate; related (test) or unrelated (control) to the target). Examples are shown in Table 9.3 (excluding filler items, which were paired with a pseudoword target). Participants were told to ignore the audio prime and judge the visual target as a word or pseudoword via button press.

All primes were monomorphemic, disyllabic nouns containing a medial singleton or geminate consonant, and were stressed on the first syllable. We matched the distribution of medial consonants among the two sets of test conditions (so the medial consonant was [p] for ten singleton test primes and [pː] for ten geminate test primes, etc.) in order to avoid any possible effects of consonant type across geminate or singleton status in participants’ responses to test items. No prime was part of a minimal pair differing only in consonant duration (i.e. shortening a medial geminate, or lengthening a medial singleton, always resulted in a non-word). In test and control items, the medial consonant did not appear in a cluster, although syllabic sonorants

TABLE 9.3. Test and control conditions

	Condition	Prime (aural)	Target (visual)
Test (related) singleton	Singleton word	[ʃtʊpə] <i>living room</i>	
	Singleton-to-geminate nonword	*[ʃtʊpːə]	
Control (unrelated) singleton	Singleton word	[ʃvɛfəl] <i>sulphur</i>	KAMMER <i>chamber</i>
	Singleton-to-geminate nonword	*[ʃvɛfːə]	
Test (related) geminate	Geminate word	[sʊpːə] <i>soup</i>	
	Geminate-to-singleton nonword	*[sʊpə]	
Control (unrelated) geminate	Geminate word	[mɛsːɪŋ] <i>brass</i>	BOUILLON <i>broth</i>
	Geminate-to-singleton nonword	*[mɛsɪŋ]	

were permitted to occur as the second, unstressed syllable (e.g. *Sattel* [sat:l] 'saddle'). All primes were selected such that manipulating medial consonant duration (described below) would result in a phonotactically acceptable word of Swiss German.

Word frequency was matched as closely as possible within each of the consonant type groups using the CELEX corpus (Baayen et al. 1995) for Standard German (as no searchable corpus of Swiss German exists). Where necessary, these rough frequency estimates were supplemented by the first author's native-speaker intuitions (native speakers' intuitions have been found to be an accurate reflection of actual frequency, Segui et al. 1982). Control primes' medial consonant differed from their test primes in place of articulation, manner of articulation, or both, and were selected to match the frequency of the test prime as closely as possible.

We recorded a female native speaker of Zurich Swiss German reading real-word primes (e.g. [sʊp:ə] 'soup' or [ʃtʊpə] 'living room') and preliminary versions of their non-word counterparts (e.g. *[sʊpə] or *[ʃtʊp:ə]). Recordings were made using an Olympus LS-11 digital recorder and Shure SM27 microphone, attached to microphone pre-amplifier model ART Studio V3; amplitudes were normalized using Praat (Boersma and Weenink 2011). Our speaker occasionally produced voicing in singleton stops, but this is not a systematic cue to phoneme identity (Brunner 1954; Kraehenmann 2001, 2003). To ensure that duration was the primary cue to consonant status in these primes, we manipulated stimuli in a modified cross-splicing procedure. The full closure duration (excluding VOT for stops) of a preliminary non-word geminate (*[ʃtʊp:ə]) was cross-spliced into the preliminary word singleton ([ʃtʊpə]) to yield a singleton-to-geminate non-word prime. To recreate a singleton real word prime (without occasional voicing), the duration of the medial consonant in the singleton-to-geminate non-word prime was shortened to match the duration of the preliminary recording. To create a geminate-to-singleton non-word prime, the closure duration of the medial consonant in the preliminary geminate word recording ([sʊp:ə]) was shortened to match the duration of the medial consonant in the preliminary non-word singleton (*[sʊpə]). Finally, we also manipulated word-geminate primes in order to avoid the factor of acoustic manipulation as a confound in responses. Word-geminate primes were created by removing two to three non-consecutive glottal pulses of the original word-geminate recording's medial consonant duration. *T*-tests were performed to ensure singleton and geminate durations were distinct, comparing real-word geminates to singletons, and singleton-to-geminate non-words to geminate-to-singleton non-words. Highly significant effects were found for both tests: for real words, $t(760) = 36.91$, $p < .0001$; for non-words, $t(760) = 32.17$, $p < .0001$.

Since targets were visually presented, and there is no orthographic convention for Swiss German, targets in this experiment were Standard German nouns. As far as possible, word targets were matched for frequency, number of letters, and number

of syllables. Semantic relatedness was judged by 13 native speakers of Swiss German (aged 17–58, average 26.7, SD: 12.46) via a web-based questionnaire. On a scale of 1 (not at all related) to 5 (extremely related), the overall rating for test pairs was 4.07 (with none rated below 3.5), and all controls were rated either 1 or 2 (average rating: 1.15).

The 50:50 ratio of semantically related to unrelated trials was maintained in the filler trials, where half of the target non-words were generated on the basis of words semantically related to the filler prime. All were phonetically and orthographically plausible words of Standard German.

9.4.2 Participants

Twenty-two subjects (all right-handed, as determined by the Edinburgh Handedness Test, Oldfield 1971; average age: 27.4 years, standard deviation: 3.8 years) were paid for their participation. All were native speakers of Swiss German from birth and were living in or visiting Oxford, England, at the time of testing. Dialectal variation could not be controlled for, but all variants of Swiss German display the robust singleton/geminate distinction (Kraehenmann 2003) tested in this experiment.

9.4.3 Procedure

Each target appeared with each of its four possible primes (test-word, test-non-word, control-word, control-non-word), making a total of 768 experimental trials. These were evenly split over four pseudo-randomized test blocks of 192 trials each, such that each target appeared exactly once in each block. Every participant saw all four experimental blocks in a randomized counterbalanced order. Each trial began with a white fixation cross which appeared in the middle of a black screen for 300 ms. 250 ms later, the auditory prime was played. After an inter-stimulus interval of 250 ms, the target was displayed for 300 ms. Participants were given up to 1,000 ms to respond; the next trial began 300 ms after their response. Reaction times were recorded from the onset of the target displayed on screen.

9.4.4 Behavioural and ERP data collection

Subjects were tested individually inside a soundproof, electromagnetically shielded booth. EEG recordings were made using a Biosemi ActiveTwo amplifier from 64 sintered Ag/AgCl pin electrodes set up in a 10/10 montage, online referenced to the mastoids. Electrooculography (EOG) recordings were made using four facial electrodes (IO1, IO2, LO1, LO2). All electrode offsets (comparable to impedance when using an active-electrode system) were kept below 30 mV; signals were digitised at 2048 Hz.

All blocks were presented to subjects using Presentation software on a 22-inch monitor about 100cm in front of the participant. Visual targets were displayed in ‘Times’ font, 100 points. Each test block was approximately 12 minutes long. Speed

and accuracy of response were equally stressed. Participants completed a 10-trial practice block before the start of recording, repeated until they felt comfortable with the procedure.

9.4.5 ERP data processing

EEG data were processed using Brain Electrical Source Analysis (BESA; MEGIS Software GmbH, Gräfelfing, Germany). Continuous EEG data recordings were filtered at a 0.01 Hz high-pass and a 30 Hz low-pass filter. Data from a pre-experimental eye movement task were processed to extract characteristic scalp topographies from eye artefacts, which were applied to the experimental recordings using an EOG correction algorithm (Ille et al. 2002). To reduce the influence of non-EEG noise, a semi-automatic procedure was applied to reject trials which exceeded an amplitude of 100 μ V or a gradient of 75 μ V.

9.4.6 Results

In both analyses, due to coding issues, data from the same block were excluded for 11 of 22 participants. Two participants were excluded: one fell asleep, one turned out to be an early L2 learner of Swiss German (rather than a native speaker from birth). Only significant results are reported in the text below.

9.4.6.1 Behavioural data Reaction times (RTs) were excluded from further analysis if they were incorrect or missed, and if they fell beyond two standard deviations of the mean RT for each subject for each condition. In addition to the data omissions above, data from twelve test items had to be excluded across all participants due to coding issues which affected the recording of behavioural responses only.

The statistical model applied, using the statistical software JMP (JMP®, Version 12. SAS Institute Inc., Cary, NC, 1989–2007), was a linear mixed-effects model with restricted maximum likelihood (REML; see Newman et al. 2012 for a similar analysis); dependent variable: reaction time, independent variables: prime type (semantically related test vs. unrelated control items), prime status (e.g. real word [ʃtʊpə] vs. non-word *[ʃtʊp:ə]) and length category ('short': singleton real words and singleton-to-geminate non-words; 'long': geminate real words and geminate-to-singleton non-words). Participants were listed as a random factor.

A summary of RT results across conditions is given in Table 9.4. Overall, responses to targets following real words were faster than following non-words, but they did not significantly differ ($F(1,5724) = 3.82, p = .051$). The effect of prime type (test vs. control) was highly significant ($F(1,5724) = 60.03, p < .0001^{***}$). A pairwise comparison established that RT to semantically related primes differed significantly from semantically unrelated control primes (rightmost row, Table 9.4), confirming the overall success of lexical access.

TABLE 9.4. Summary of RT results

Condition		Example of prime	Example of target	Mean RT in ms (standard error)	Priming effect in ms	sig.
Singleton word	Test (related)	[ʃtʊpə] <i>living room</i>	KAMMER <i>chamber</i>	681 (20.83)	17	$p = .004^{**}$
	Control (unrelated)	[ʃvefəl] <i>sulphur</i>		698 (20.84)		
Singleton-to-geminate prime	Test (related)	*[ʃtʊp:ə]		682 (20.83)	26	$p < .0001^{***}$
	Control (unrelated)	*[ʃvef:əl]		708 (20.85)		
Geminate word	Test (related)	[sʊp:ə] <i>soup</i>	BOUILLON <i>broth</i>	669 (20.82)	29	$p < .0001^{***}$
	Control (unrelated)	[mes:ɪŋ] <i>brass</i>		698 (20.83)		
Geminate-to-singleton prime	Test (related)	*[sʊpə]		679 (20.82)	20	$p = .001^{**}$
	Control (unrelated)	*[mesɪŋ]		700 (20.82)		

Effects on response accuracy were analysed using a logit Generalized Linear Model (treating response accuracy as a binomial distribution) for the fixed effects of prime type, word status, and length category. In the analysis of response accuracy, there were significant effects of prime type, $\chi^2 = 7.70$, $p = .006$ and length, $\chi^2 = 36.60$, $p < .0001$, and an interaction of prime type and word status, $\chi^2 = 5.87$, $p = .015$. Difference-of-differences planned comparisons were performed on the non-significant three-way interaction, comparing the difference between control and test items between words and non-words for both long and short durations. These showed a significant priming difference in the ‘long’ category (geminate real word and geminate-to-singleton non-word), $\chi^2 = 5.80$, $p = .016$, but not in the ‘short’ category (singleton real word and singleton-to-geminate non-word), $\chi^2 = 0.68$, $p = .411$. Participants were thus significantly more prone to response errors following geminate-to-singleton than following singleton-to-geminate manipulated primes, suggesting that geminate-to-singleton manipulation had a greater disruptive effect on lexical access than singleton-to-geminate manipulation.

9.4.6.2 ERP data EEG epochs were averaged from target stimulus onset for all correct responses to word targets within a pre-stimulus baseline period of 200 ms and a window of 1000 ms from target onset. N400 effects are particularly robust in

the centro-parietal region (Kutas and Hillyard 1984; Domalski et al. 2013). Therefore, the region of interest (ROI) selected for further analysis was a group of nine electrodes clustered around and including Pz: CP1, CPz, CP2; P1, Pz, P2; PO3, POz, PO3. In the control conditions, the N400 peak was observed most clearly around 450 ms after target onset. For further analysis we selected a 200 ms window centred around 450 ms after target onset (see Proverbio et al. 2004; Filik and Leuthold 2008).

The figures which follow show grand average ERPs for test vs. control within each of the experimental categories averaged across electrodes selected for the ROI: singleton real words (Figure 9.3a), geminate real words (Figure 9.3b), singleton-to-geminate non-words (Figure 9.3c), and geminate-to-singleton non-words (Figure 9.3d). The 350–550 ms time window selected for later ERP amplitude analysis is highlighted in each figure.

As in the analysis of behavioural results, N400 amplitudes were analysed with a linear mixed-effects model with restricted maximum likelihood (REML); dependent variable: mean ERP amplitude in microvolts (extracted for the time window 350–550 ms after target onset), independent variables: prime type (semantically related test vs. unrelated control items), prime status (e.g. real word [ʃtʊpə] vs. non-word *[ʃtʊp:ə]) and length category ('short': singleton real words and singleton-to-geminate non-words; 'long': geminate real words and geminate-to-singleton non-words). Participants were listed as a random factor.

An overview of the ERP results is provided in Table 9.5. The effect of test vs. control items was highly significant ($F(1,1413) = 78.8$; $p < .0001^{***}$). Differences in N400s to targets following real word and non-word primes were not significant ($F(1,1413) = 2.70$; $p = 0.101$). The linear mixed effects model also revealed a significant three-way interaction ($F(1,1413) = 5.43$; $p = .20^*$) between test/control primes by real word/non-word primes by categories. A pairwise comparison established that N400 responses differed significantly between semantically related test and unrelated control items in all conditions, i.e. priming was successful where primes and targets were semantically related, in spite of the duration manipulations.

However, the magnitude of the priming effect differed between non-word prime types. A difference of differences contrast test (reported in Table 9.6) compared the difference between the N400s of singleton real word vs. singleton-to-geminate non-word with the difference between N400s of geminate real word vs. geminate-to-singleton non-words. This test established that while there was no significant difference between the N400 amplitudes of singleton real word vs. singleton-to-geminate non-words, there was a significant difference between geminate real word vs. geminate-to-singleton non-words. Shortening geminates into singletons therefore had a stronger adverse effect on lexical access than lengthening singletons into geminates.

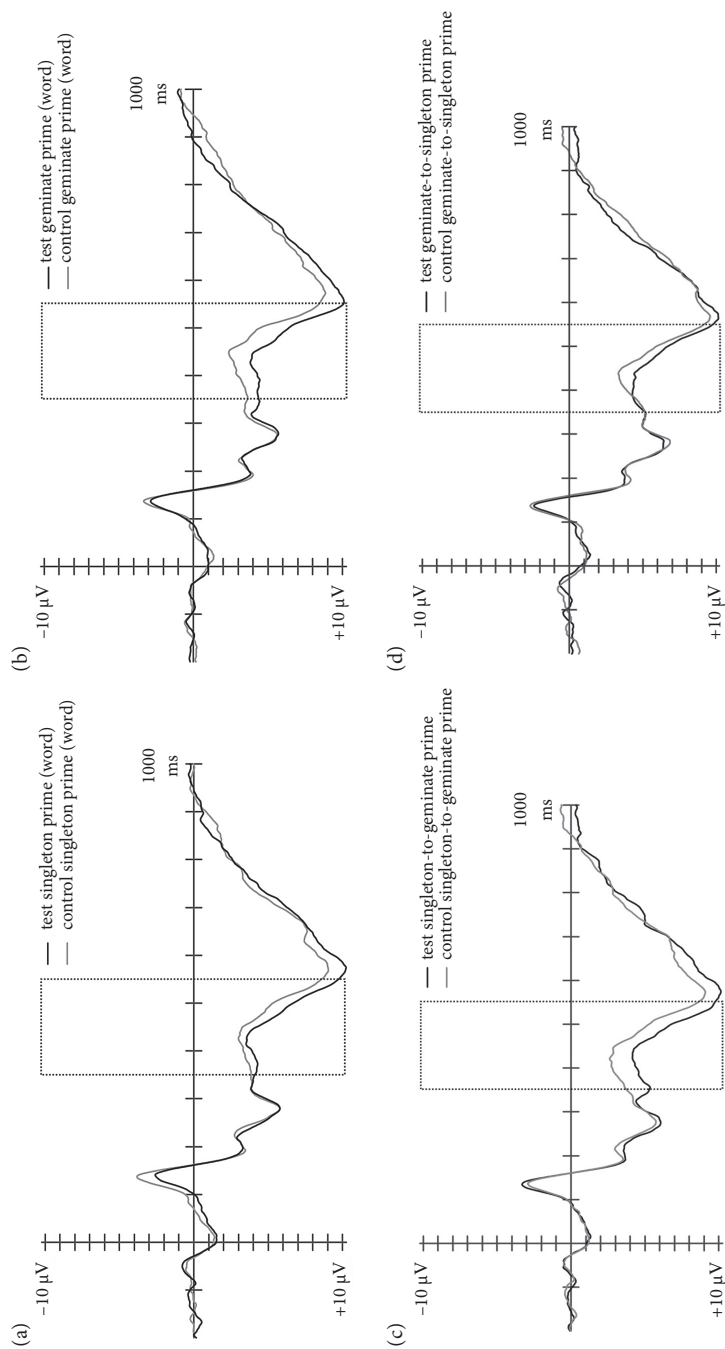


FIGURE 9.3 (a) Averaged ERP responses to singleton real words. (b) Averaged ERP responses to geminate real words. (c) Averaged ERP responses to singleton-to-geminate non-words. (d) Averaged ERP responses to geminate-to-singleton non-words.

TABLE 9.5 Summary of ERP results

Condition		Example of prime	Example of target	Mean ERP amplitude in μ V (SE)	Priming effect in μ V (SE)	Sig.
Singleton word	Test (related)	[ʃʊpə] <i>living room</i>	KAMMER <i>chamber</i>	6.35 (0.721)	-0.97 (0.228)	$p < .001^{***}$
	Control (unrelated)	[ʃvɛlə] <i>sulphur</i>		5.38 (0.721)		
Singleton-to-geminate prime	Test (related)	*[ʃʊp:ə]		6.49 (0.721)	-1.33 (0.228)	$p < .001^{***}$
	Control (unrelated)	*[ʃvɛf:əl]		5.17 (0.721)		
Geminate word	Test (related)	[sʊp:ə] <i>soup</i>	BOUILLON <i>broth</i>	6.37 (0.721)	-1.23 (0.228)	$p < .001^{***}$
	Control (unrelated)	[mes:ɪn] <i>brass</i>		5.13 (0.721)		
Geminate-to-singleton prime	Test (related)	*[sʊpə]		6.43 (0.721)	-0.53 (0.228)	$p = .021^*$
	Control (unrelated)	*[mesɪn]				

TABLE 9.6. Difference-of-difference test for ERP amplitudes

Test vs. control in geminate real words vs. geminate-to-singleton non-words	$t(1413) = -2.19$	$p = .029^*$
Test vs. control in singleton real words vs. singleton-to-geminate non-words	$t(1413) = 1.11$	$p = .267$ <i>n.s.</i>

9.5 Discussion

This study targeted the interaction of processing at the level of melody and timing in geminate recognition, with the overall aim of elucidating how consonant duration in the acoustic signal is mapped onto underlying phoneme representations, and what this implies for the representation of geminates in the mental lexicon. We aimed to test whether Swiss German, which (unlike Bengali) has the representational capacity to encode variation in duration in all positions within the word, would nonetheless display an asymmetry in processing geminate duration. Our results confirm this hypothesis, suggesting that the Swiss German geminate/singleton contrast is underspecified in the mental lexicon. Our results also confirm that the asymmetry observed in Bengali is due to a geminate representation which includes the structural level, not a lack of representational capacity for encoding the geminate/singleton contrast other than in word-medial position.

We concentrate our discussion on two main experimental effects. First, the analyses of N400 amplitude and reaction times revealed a reduction of the priming effect in response to geminate-to-singleton non-words, but not singleton-to-geminate non-words. In accounting for these results, we explore representation-based and processing-based sources for the effects we observe. We suggest that this processing asymmetry is evidence of a structurally underspecified representation of the geminate/singleton contrast, and discuss what form the structural specification might take. We also speculate about a processing-based alternative: that this asymmetry reflects a non-linear speech parsing process (‘CV template parsing’) by which speakers preferentially interpret consonants in the acoustic signal as onsets.

Secondly, although manipulating consonant duration reduced priming, it did not, overall, cause lexical access to fail (unlike in similar experiments in Bengali, Roberts et al. 2014; Kotzor et al. 2015). We suggest that this is due to a combination of neighbourhood cohort effects and Swiss German participants’ facility in switching between native and non-native (Standard German) phonologies. Repetition effects may also be a factor.

9.5.1 *Representational account: a structurally underspecified geminate/singleton contrast*

Although listeners are successful overall in recovering from variation in duration, processing of Swiss German geminate duration nonetheless displays the telltale asymmetry that suggests an underspecified mental representation. Thus only acoustic information that meets the criteria for geminate recognition can be mapped onto an underlying geminate, whereas consonant information of any duration can be mapped onto an underlying singleton. However, it is unclear what form this underspecified representation should take.

Existing accounts of underspecification operate at the melodic level: for instance, if German stops are specified for [LABIAL], consonants in the input can only be mapped onto an underlyingly [LABIAL] consonant if the parser can extract acoustic evidence to satisfy the requirements of that [LABIAL] feature (Lahiri and Reetz 2002, 2010). In theory, geminates could be represented as specified at the melodic level, e.g. through a privative feature [LONG]. This option can be excluded on external logical and empirical grounds. First, geminate duration is not a feature, for a host of theoretical reasons (Lahiri and Marslen-Wilson 1992; Kraehenmann 2003; see also section 9.3). Secondly, if the representation of the geminate/singleton contrast were a feature [LONG] ~ Ø, we would expect very different experimental results in Lahiri and Marslen-Wilson (1992): listeners should have been able to map long consonant duration onto an underlying geminate as soon as they had enough evidence in the signal to confirm that [LONG] was present. Yet even when given the consonant's full closure duration, they interpreted the consonant as a singleton. However, evidence of syllable structure (in the form of sonorant release and two glottal pulses of a following vowel) allowed participants to identify geminates above chance, suggesting that geminates are specified at a structural, not melodic, level in Bengali.

Given the results of our processing experiment, some kind of asymmetric specification is clearly also required for Swiss German geminates. At its most fundamental, the specification must be structural, rather than featural. There are three candidates for such a structural unit: a CV-skeleton (McCarthy 1982; Clements and Keyser 1983), an X-skeleton in which the difference between consonants and vowels is neutralized (Levin 1985), and a moraic representation which doubles up as a weight-bearing unit. All three options assume (and indeed this is universal) that a geminate is intrinsically heterosyllabic. If it happens to be found at the edge of a word, the geminate will associate with the coda; if it is word-initial, it is structurally indistinguishable and acoustically neutralized, but its presence is felt in articulation (Kraehenmann and Lahiri 2008). Note that the representation of the geminate/singleton contrast within both the CV- and X-skeleton theories is binary (singletons are linked to one timing slot, X, whereas geminates are doubly linked, XX) rather than privative, making both of these representations impossible to reconcile with our results in processing. Since our results point to an asymmetric representation, the only real candidate is the mora,

as Kotzor et al. (Chapter 8) also propose for Bengali. However, at this present juncture, given Kraehenmann's (2003) account that Swiss German geminates do not necessarily add to weight (see also Mohanan and Mohanan 1984 for Malayalam, and Lahiri and Koreman 1988 for Dutch), we would like to remain agnostic about the precise nature of the representation. What is clear, however, is that whatever is encoded in the lexicon to mark the geminate/singleton contrast has to reflect the fact that there is an asymmetry and that medial geminates have to be heterosyllabic.

Adapting this structural specification to a FUL-style model, only phonetically long consonant duration satisfies the requirement of this structural unit so as to be mapped onto an underlying geminate ([MATCH] response). A phonetically short consonant will therefore fail to be mapped onto an underlying geminate ([MISMATCH]), whereas a phonetically long consonant can be mapped onto an underlying singleton that is structurally underspecified ([NO-MISMATCH]). The predictions of this model exactly match our results. Geminate perception in Swiss German as well as Bengali therefore confirms that the radical underspecification account of the mental lexicon can and should be extended beyond featural information into the structural level of representation.

If geminates have an underlying structural specification that causes them to be syllabified into two syllables, and if only a phonetically long duration can be mapped onto a geminate with this underlying structural specification, this predicts that geminate recognition will encounter difficulties everywhere except in medial position. In running speech, word-initial and final geminates resyllabify with the coda of a preceding and the onset of a following syllable, respectively, and therefore usually appear as heterosyllabic, so recognition should not be hampered here. While the syllabification and representation of utterance-initial and final geminates is controversial (Kraehenmann 2003), it is safe to assume that they cannot be heterosyllabic. Interestingly, Swiss German speakers confirm the prediction that geminate perception will suffer at syllable edges: they are unable to distinguish geminates from singletons utterance-initially (Kraehenmann 2001; Kraehenmann and Lahiri 2008). To our knowledge, the perception of utterance-final geminates has not been investigated, but if geminates indeed bear a structural specification which requires them to be syllabified across two syllables, this would predict difficulty in the perception of utterance-final geminates. Cross-linguistically, final geminates are rare and initial geminates even more so (Kraehenmann 2011); difficulty in recognition may be a part of why edge geminates are typologically marked.

9.5.2 *Processing account: a CV template in parsing*

An alternative approach is to explore the possible role of parsing processes in yielding the asymmetric results we observe. We speculatively propose a mechanism at the

syllable level of speech recognition: the CV template. In a first pass of processing, incoming acoustic information is roughly chunked left to right into consonants and vowels on the basis of melodic information only and, as far as possible, slotted into a simple CV syllable template.³ In this way, all consonants, regardless of duration, are initially interpreted as the onset of a CV syllable, and lexical access to a word containing a medial singleton can proceed. In geminate recognition, the initial parse into CV syllables must be revised in a second pass that includes evaluation of durational information, such that the melodic information extracted from the signal can ultimately be mapped onto two syllables. ERP measures may reflect the effects of such a second pass (such as a processing latency, which is likely to be erased by long-distance processing measures like reaction times), but given the design and statistical power of our experiment, we were not able to draw any conclusions about this, and leave it to future research to clarify.

If the CV template operates in speech recognition, there must be some mechanism that applies asymmetrically in prompting the second pass for geminate recognition. It cannot be phonetic consonant duration itself: if so, Swiss German listeners would have integrated the duration of singleton-to-geminate consonants, but contrary to this prediction, priming was reduced in this experimental condition. It is possible that melodic information is, by and large, sufficient to activate a cohort of candidates, including one word candidate containing a geminate which prompts evaluation of consonant duration, resyllabification (if phonetic duration is indeed long), and ultimately, lexical access. This account is troublesome for geminate/singleton minimal pairs, however, since the melodic information for the two words is identical, and the word containing a singleton would always be activated more. Identification of minimal pairs with initial geminates or singletons is indeed problematic (Kraehenmann 2003; Kraehenmann and Lahiri 2008), but no such issues are reported for medial geminate/singleton minimal pairs. How a geminate would prompt evaluation of consonant duration needs further discussion: for instance, this might rely on a representational factor, e.g. a structural specification on the geminate, such that parsing and representation might complement each other in giving rise to the phenomena of geminate processing.

The CV template could explain why Swiss German speakers have difficulty perceiving utterance-initial geminates (Kraehenmann and Lahiri 2008): if there is no preceding syllable to resyllabify a consonant with, yet geminate recognition relies on a second pass of resyllabification, recognition of a segment as a geminate would fail. On the other hand, ignoring duration except where required might confer an advantage in processing a variable speech signal. Cross-linguistically, prosodic

³ This cannot apply when the acoustic input contains a sequence of consonants. We tentatively suggest that the first might form the head of a defective template which is then syllabified with another template according to language-specific phonotactic rules.

factors like speech rate (Pickett et al. 1999; Hirata and Whiton 2005) and stress (Payne 2005; Payne and Eftychiou 2006) can lengthen absolute consonant duration. If duration were always evaluated (instead of, as we propose, only being evaluated when triggered by top-down information, whatever its nature), this could introduce a potential for misanalysis into speech processing. However, if parsing is guided by a CV template augmented by a geminate representation which causes durational information to be evaluated, prosodically lengthened consonants would never be confused with underlying geminates. Where postlexical assimilation processes result in concatenated geminates, two underlying segments are realized as a single phonetically long consonant. How this would be mapped onto two underlying segments within the framework of a CV template is problematic, but non-lexical geminates are beyond the scope of this chapter, and we leave this to be explored in future work.

Whether the phenomena of geminate processing are due to geminates' representation in the mental lexicon or are a product of parsing routines (which, as we discuss, are not mutually exclusive possibilities), it seems clear that models of lexical access must incorporate a structural level of processing (beyond the melodic level). Existing models, such as TRACE (McClelland and Elman 1986) or SHORTLIST (Norris 1994), rely on melodic information alone in word recognition. Yet access to structural information is required to explain the asymmetry in geminate processing found in Swiss German and in Bengali (Roberts et al. 2014; Kotzor et al. 2015; Kotzor et al., Chapter 8). If the CV template is available in languages with a geminate/singleton contrast, it may also be operating behind the scenes in languages with no such contrast. For instance, even in English, it seems that some syllable-level phenomena, e.g. stress, can disrupt the otherwise more left-to-right parse of segments as a sequence (Cutler and Norris 1988).

Overall, although a CV template that guides word recognition could explain the syllable-level effects we observed, further research is required to determine what asymmetric mechanism triggers evaluation of durational information: whether the geminate/singleton contrast is structurally underspecified in the mental lexicon, a CV template guides word recognition, or a combination of both.

9.5.3 *Swiss German speakers integrate durational variation*

Our results show a reduced priming effect in response to the geminate-to-singleton related condition; but nonetheless there was still a significant effect of relatedness in the geminate-to-singleton related vs. unrelated conditions, such that lexical access was hindered but not disrupted. That is, Swiss German speakers tolerate variation in consonant duration, though less when a geminate is shortened than when a singleton is lengthened. This may be due to the availability of acoustic information beyond the medial consonant's duration, and the use of non-Swiss German phonological systems to ignore conflicting durational information.

Lexical recognition can take place before the entirety of the information present in the acoustic signal has been evaluated (see Marslen-Wilson and Zwitserlood 1989). In this experiment, listeners had access to all of the acoustic information after the medial consonant, as well as that consonant's melodic information, helping to narrow down the set of possible candidates. Since none of the primes used in this experiment belonged to minimal pairs differing only in duration, a neighbourhood cohort would have been sufficient to select the appropriate word candidate, even without access to the relevant structural information that would have identified the consonant as a geminate or singleton. If, indeed, duration evaluation is only initiated through a structural specification or a second pass in syllable parsing, consonant duration would be ignored in most of the primes participants heard. Why this is possible in Swiss German but not in Bengali (Chapter 8) may be a result of Swiss Germans' high degree of multilingualism and close contact with non-native phonologies.

As multilinguals, our participants routinely access the phonologies of English and Standard German, which do not contrast their consonant series on the basis of phonological quantity, meaning that consonant duration is never exploited when participants process second-language input. Bear in mind that participants were responding to visual targets in Standard German (since Swiss German has no written standard, see section 9.2.2), which may have encouraged them to access the phonological system of Standard German. Past research has shown that Swiss German speakers are highly skilled in exploiting non-native contrasts: Pohl (2011) demonstrates that Swiss German participants can exploit Standard German's phonemic durational variation in VOT to correctly categorize stops, even though VOT is not a distinctive feature in the phonological system of Swiss German. On the other hand, Standard German speakers presented with stimuli with constant VOT but variable consonant duration displayed only a tendency to map geminate and singleton stops onto the correct Standard German stop categories. Swiss German speakers' extensive experience with the phonology of Standard German may therefore have allowed them to ignore durational variation in the present study's stimuli.

As one thoughtful reviewer pointed out, our results are likely to have been strongly affected by repetition effects. We had to balance the demands of statistical power with an extremely sparse participant population, and therefore, each participant saw each word target four times, combined once with each prime condition (test, control, real word, and non-word). Repeating targets tends to reduce priming effects overall (Grill-Spector et al. 2006), reflected in the symmetric response seen in the RT data. The fact that the non-word primes asymmetrically reduced priming in the geminate-to-singleton condition despite repetition of targets suggests that the processing asymmetry is quite robust in ERPs in Swiss German—though perhaps, for the reasons discussed above, somewhat more fragile than in Bengali. A repetition of the behavioural portion of this experiment using a traditional Latin square design and more participants is required to confirm our asymmetric result.

9.5.4 *Open questions*

Finally, we briefly note an avenue for further research. Our focus on the structural (rather than melodic) level of representation for geminates is driven by theoretical concerns that rule out representing length as a feature, but also by the experimental results of Lahiri and Marslen-Wilson (1992), in which participants only identified geminate consonants above chance when, in addition to the full closure duration, they also heard the release and the first two glottal pulses of the following vowel. The authors interpreted this as evidence that structural information is required before a consonant can be identified as heterosyllabic and therefore, a geminate. We note, however, that in the case of medial geminates, consonant duration and heterosyllabicity are confounded. Future exploration must therefore determine to what extent geminate recognition is indeed crucially dependent on the availability of syllable-level information. This could be achieved by investigating the perception of contrastive vowel duration (e.g. in Japanese) or the perception of final consonant duration in Swiss German geminates and singletons. In both cases, whether the segment is long or short, there is no difference in syllabification, i.e. duration is decoupled from structural information.

9.6 Conclusion

Geminate consonants offer a unique window into the phonetics–phonology interface in speech recognition: a single acoustic cue, consonant duration, determines both the phoneme’s identity and the word’s syllable structure. The present study investigated geminate representation through the lens of these interacting factors in the processing of medial geminates in Swiss German. ERP and behavioural measures of the success of lexical access to primes with manipulated medial consonant duration revealed a significant processing asymmetry. Listeners were able to map a too-long consonant onto an underlying singleton, but not a too-short consonant onto an underlying geminate, suggesting that consonant duration is not evaluated unless required by top-down information. We argue that a FUL-style conception of the mental lexicon (Lahiri and Reetz 2002, 2010) must include a structural level of underspecification, such that singletons are underspecified in the mental lexicon while geminates are specified with a structural unit that asymmetrically triggers the evaluation of phonetic consonant duration in the signal. The processing asymmetry may alternatively be explained as the result of a speech parsing routine, the CV template, which assigns all consonant information in the signal to a syllable onset in the first pass of processing, only evaluating consonant duration and reanalysing syllable structure when prompted by top-down information. More research is required to identify whether geminates are processed asymmetrically because of the constraints of representation, processing, or both, as well as to further pinpoint the role of the syllable in lexical access.

Acknowledgements

This chapter is based on the first author's MPhil thesis (Oxford, 2013). Earlier versions were presented at the 2014 Annual Meeting of the Linguistics Association of Great Britain in Oxford and at the satellite workshop on geminate consonants across the world at ICPHS 2015. We are indebted to our friends and colleagues, Naomi Feldman, Bill Idsardi, Haruo Kubozono, Andrew Nevins, Colin Phillips, Henning Reetz, and Philip Roberts, for invaluable discussion on early incarnations of this work. We are also grateful for the thoughtful comments we received from Pierre Hallé and one anonymous reviewer. Finally, we thank Ann Ehrenhofer-Murray for lending her voice to this experiment, and our participants for their patience. This work was supported by a Fulbright Postgraduate Student Award, a University of Maryland Flagship Fellowship, a Sloane Robinson Graduate Award, and a Lincoln College, Oxford, research grant to Lara Ehrenhofer; and European Research Council grant FP7-IST-269670 (PI: Aditi Lahiri).