Asymmetries in the perception of non-native consonantal and vocalic length contrasts

Abstract

How well can non-native length contrasts for vowels and for consonants be perceived and is one more difficult than the other? Three listener groups (native Italian and German as well as advanced German learners of Italian) performed a speeded same-different task involving vocalic and consonantal length contrasts as well as segmental contrasts as controls. Phonologically, Italian, but not German, has a consonantal length contrast, while German, but not Italian, has a vocalic length contrast. Analysis of responses yielded a clear asymmetry: A non-native *vowel* length contrast was perceived just as well as the native consonantal length contrast, however, a non-native *consonantal* length contrast was perceived poorly compared to the native vocalic length contrast: Italians showed higher sensitivity for consonantal length than German learners of Italian, who in turn were better than German non-learners. Reaction time analyses indicated that, despite being more successful, the decision was just as difficult for learners as for non-learners, suggesting different types of difficulty for listeners with and without experience with a consonantal length contrast.

Keywords: vocalic length, consonantal length, Italian, German, advanced learners, second language phonology, L2 suprasegmentals

Introduction

In the course of adult second language (L2) acquisition, the L2 sound system remains to be a big challenge for the vast majority of learners even after other grammatical (e.g., morphosyntactic) facets of the target language have been mastered. This usually results in a distinct and discernable non-native accent (cf. Ioup, 1984; Munro and Derwing, 1995), which has been shown to slow down speech comprehension in native listeners (Braun et al., 2011a, 2011b) and may even lead to lexical confusion. Dutch listeners of English, for instance, find it very hard to distinguish between English *cattle* and *kettle,* as the Dutch phoneme inventory only contains one half-open front vowel /ɛ/, instead of two as in English: /ɛ/ and /æ/ (Escudero et al., 2008; Weber and Cutler, 2004).

In order to address such persistent L2 learner problems, the focus of research in the area of second language acquisition since the 1980s has developed from broad pedagogical or theoretical models, which were usually based on morphosyntactic learner productions (e.g., Monitor Model including the Input Hypothesis, cf. Krashen, 1985, or analyses couched in UG framework, cf. White, 1989, 2003), towards research aiming at more specific investigation of the different linguistic subfields. One subfield that yielded a large body of psycholinguistic studies has been segmental interlanguage phonology (for an overview see Altmann and Kabak, 2011). This, in turn, resulted in the development of a number of theories regarding L2 speech perception both phonetically oriented (e.g. Perceptual Assimilation Model, cf. Best, 1995; Speech Learning Model, cf. Flege, 1995; Native Language Magnet, cf. Kuhl, 1993), or phonological (e.g. Phonological Interference Model, cf. Brown, 1998, 2000; Ontogeny-Phylogeny Model, cf. Major, 2002). Furthermore, mainly over the last decade, there has been increasing experimental interest in the perception and production of suprasegmentals such as stress (e.g. Altmann, 2006; Dupoux et al., 2008; Kijak, 2009; Tremblay, 2008), tone (e.g. Chiao et al., 2011; Gandour, 1983; Wang et al., 2003), or segmental length (e.g. Hayes-Harb, 2005; Hisagi et al., 2010; McAllister et al., 2002).

The current paper reports results from a perception study on segmental length contrasts which are phonemic in the non-native language (L2) but not in the native language (L1) and compares it to the perception of L1 length contrasts. In order to provide a holistic picture, the perception of both vowel and consonant length is investigated here with identical participant groups. This procedure allows for a reliable cross-linguistic comparison of the ability to perceive non-native length contrasts for different types of segments. In addition, possible effects of experience with the L2 (as reported in Heeren and Schouten, 2008, 2010) are traced by comparing naïve L1 speakers with no experience with the target language (non-learners, as a baseline) to advanced learners of that L2.

In the following, we will first give an overview of the notion of segmental length as well as consonantal and vocalic length in L2 acquisition, followed by a review of studies suggesting improvement in the ability to perceive L2 length contrasts over time. We will then compare segmental length in Italian and German, our target languages tested, before we present the experimental procedure and results.

Background

Segmental length contrasts vs. allophonic variation

Length is a suprasegmental feature that occurs with consonants as well as with vowels in the world's languages (Ladefoged and Maddieson, 1996). Segmental length is often unpredictable and signals lexical contrasts (e.g. the distinction between /fato/ 'fate' vs. /fatːo/ 'fact' in Italian or between /ʃtat/ 'city' vs. /ʃtaːt/ 'state' in German). Consonantal length contrasts are found, among others, in Italian, Japanese, Arabic, Finnish, or Swiss German, vocalic length contrasts in Japanese, Swedish, Standard German, or Finnish.

Segments may however also appear predictably lengthened or shortened based on their position in the word or utterance, over and above durational variation due to differences in speech rate. On a local level, vowel duration is dependent on the vowel in question (high vowels being intrinsically shorter than low vowels, cf. House and Fairbanks, 1953) and also on the immediate segmental context (e.g. longer before voiced consonants than before voiceless ones, cf. Peterson and Lehiste, 1960, and longer before fricatives than before plosives or affricates, cf. Delattre, 1966).

More globally, phonemes are lengthened in phrase-initial (Cho and McQueen, 2005; Fougeron and Keating, 1997; Wightman et al.,1992) and phrase-final position (Turk & Shattuck-Hufnagel, 2007; Wightman et al., 1992). A further reliable effect is polysyllabic shortening, i.e., the longer a word, the shorter its syllables and segments (e.g. Braun and Geiselmann, 2011; Lehiste, 1972; Port 1981; White and Turk, 2010). Duration is further influenced by word stress and sentence accent: vowels are lengthened in stressed syllables but shortened in unstressed syllables (e.g. Baumann et al., 2006; Braun et al., 2011b). Such allophonic changes in duration are also found for consonants. For instance, when the same phoneme occurs at the end of one word or morpheme and again at the onset of the next word or morpheme (as in German *mit Tim* ‘with Tim’ or *mitteilen* 'to convey'), a 'quasi geminate' is produced, which is longer than an otherwise identical single consonant (Mikuteit, 2007). Consonant duration is also increased at the start of prosodic phrases (phrase-initial strengthening, cf. Fougeron and Keating, 1997; Cho and McQueen, 2005). It is well documented that such non-phonemic durational variation is used in online speech comprehension (e.g. Salverda et al., 2003; Tagliapietra and McQueen, 2010).

Phonetically, contrastive length features as segment duration. Psychoacoustic studies have shown that participants' sensitivity to detect general durational differences increases as the duration of the stimuli increases (e.g. Abel, 1972;

Creelman, 1962). For a stimulus duration of 100ms, for instance, the minimal detectable duration difference is 15 ms (Abel, 1972). This indicates a high perceptual sensitivity for durational contrasts in non-linguistic auditory input. Such unbiased high sensitivity for stimulus duration is, however, not upheld in language contexts. When processing linguistic speech stimuli, the perception of duration is influenced by the native phonological system, an effect that emerges early in development (e.g., Pohl and Grijzenhout, 2010, for differential discrimination of geminate and singleton stops by Swiss German and Standard German 10-12 month olds).

In languages that employ length contrasts, prelexical and lexical representations need to be tuned to durational contrasts so that they can activate the appropriate lexical candidates (McQueen, 2005). On the contrary, when length distinctions in the L1 are absent, perceptual sensitivity to duration is reduced (Gottfried and Beddor, 1988). It should be stressed, however, that perceptual sensitivity to duration is needed in all languages as segment durations do not only signal lexical contrasts but may also convey information about prosodic structure (as illustrated in detail above). In the end, only the combination of lexical and prosodic information allows us to efficiently decode the utterance interpretation.

Length in L2 acquisition

Non-target-like productions of L2 length for different classes of segments are well documented in the literature, even by advanced L2 learners (Han, 1992; Kabak et al., 2011; Mah and Archibald, 2003; McAllister et al., 2002). One explanation for such production problems may lie in non-target-like perception of length information in the L2 input, which is the main focus of the current paper.

There have been a number of studies on the perception of vowel length distinctions for L2 learners, whose L1 lacks such a contrast, employing both behavioral (e.g. Cebrian, 2006; Flege et al., 1997; McAllister et al., 2002; Ylinen et al., 2005) as well as electrophysiological methods (e.g. Hisagi et al., 2010; Nenonen et al., 2003, 2005). On the other hand, substantially less work has been done on the L2 perception of geminate vs. singleton consonants (cf. Hardison and Saigo, 2010; Hayes, 2002; Hayes-Harb, 2005; Hayes-Harb and Masuda, 2008; Heeren and Schouten, 2008). The following section gives a more detailed overview of studies concerned with the L2 acquisition of segmental length.

Vocalic length. The cross-linguistic perception of vocalic length supposedly poses little problems irrespective of the use of vocalic length contrasts in one's native language (e.g., Bohn, 1995; Cebrian, 2006; Flege et al., 1997; García Lecumberri and Cenoz, 1997). Turkish late learners, for example, discriminated German long and short vowels just as well as German native controls (Nimz, 2011). In addition, it has been found that L2 learners whose L1 does not contrast vowel length may still use duration to distinguish vowel pairs in the second language, even in cases where native listeners of that language themselves make more use of spectral differences (e.g., Bohn, 1995; Cebrian, 2006; Flege et al., 1997; Kondaurova and Francis, 2008; Wang and Munro, 1999). For instance, Spanish and Mandarin learners of English relied more on the temporal cues present in the signal for the English /i/ - /ɪ/ contrast, while L1 controls relied more on spectral information. Based on such results, Bohn (1995) argued for a "general speech perception strategy" (p. 300), which claims that listeners utilize durational cues whenever spectral segmental information is difficult to process. The explanation for such 'inappropriate' use of vowel duration is that durational cues are more salient and hence easier to perceive. Additional experimental evidence in Jacquemot et al. (2003) indicates that beginning learners may not process novel vowel duration contrasts on a phonological basis but rather as purely auditory stimuli on an acoustic basis.

Nevertheless, there have also been empirical findings relativizing this general positive view of durational cues for vowel perception. In summary, they report significantly better results for listeners whose L1 contains both long and short vowels compared to listeners with no length contrasts in their L1. McAllister et al. (2002) conducted a cross-linguistic identification task with English, Estonian and Spanish learners of Swedish as well as Swedish controls. Their results showed that Spanish learners of Swedish (no L1 length contrast) yielded significantly lower correctness scores in identifying Swedish long and short vowels than Estonian learners (L1 length contrast) or Swedish natives. However, English learners of Swedish performed significantly better than the Spanish participants (yet still lower than Swedish or Estonian groups), which was taken as potential evidence for some durational sensitivity in the English participants due to their L1 experience with allophonic vowel duration differences. Nenonen et al. (2003) conducted an electrophysiological mismatch negativity study (MMN) and compared the perception of length contrasts for early Russian L2 learners of Finnish compared to Finnish L1 listeners. They were tested on vowels that could be mapped onto Russian phonemes and those that were too different to be mapped. The Russian learners of Finnish, whose L1 does not contrast vowel length, showed a lower mismatch negativity (decreased detection of the change) than Finnish natives, but only in those sounds that could be mapped onto Russian vowels. For vowels that could not be categorized in the L1 system, sensitivity to length contrasts was not different from Finnish native listeners'. These findings suggest that, on the one hand, the effect of the L1 phonological system may be positive, in that allophonic experience with vocalic length in the L1 possibly provides some advantage in L2 speech perception. On the other hand, the L1 phonological system may affect L2 perception of vowel length negatively, in that the possibility of mapping non-native sounds onto close L1 categories blocks the application of more general temporal discrimination strategies.

Consonantal length. Regarding the non-native perception of consonantal length, comparatively less research has been reported and most of it focuses on English learners of Japanese. Hayes (2002), for instance, tested the perception of the Japanese singletons /k/, /t/ and /s/ and their corresponding geminates by different groups of English learners of Japanese as well as by Japanese natives in a same-different task.[1] Participants had to determine whether two nonwords that only differed in the length of the medial consonant were the same or different. Not surprisingly, her results showed that accuracy improved with proficiency but did not reach native-like performance. Hayes-Harb (2005) reports an identification task requiring participants to choose between 'single' or 'double' consonant, in which identical consonants were manipulated in 13 durational steps. Native Japanese speakers showed categorical perception with a clear perceptual boundary (as evidenced by a sigmoid response function) while the responses of English participants with no exposure to Japanese (and thus no experience with consonantal length contrasts) yielded a near linear response function starting at only around 80% singleton judgments for the shortest duration and ending with still at as much as about 20% singleton responses for the longest duration.[2] Results from English learners of Japanese were found to fall in-

[1] Unfortunately, no specification of the inter-stimulus-interval is provided in this study.
[2] It is difficult to assess how non-learners (whose L1 lacks a consonantal length contrast) can achieve the task of distinguishing 'single' and 'double' consonants.

between the naïve English and the native Japanese groups. Compared to the native English group, their judgments started at a higher proportion of singleton responses (around 90%) for the shortest stimuli and approximated a lower proportion of singleton judgments (around 10%). Hence, consonantal length appears to be difficult to identify, even after prolongued exposure.

In sum, there is experimental evidence indicating that the native phonological system interferes with the perception of segmental length. While durational cues may potentially be universally available to all listeners (Bohn 1995), L1 properties have been argued to affect the perception of vowel length – both in the positive (cf. McAllister et al., 2002) and in the negative sense (cf. Nenonen et al., 2003). Regarding the perception of L2 consonantal length, there is no diversity of cross-linguistic studies (or of a larger variety of segments) available. However, the existing literature at this point suggests that it is difficult to perceive an L2 singleton-geminate contrast if it is absent in the L1.

Perceptual improvement with increased exposure

In most of the studies mentioned above, even when advanced learners did not reach target-like scores, learners at a more advanced stage still outperformed non-learners and/or lower-level L2 learners from the same L1 background (e.g., Hayes, 2002; Hayes-Harb, 2005; Hardison and Saigo, 2010). From an acquisitional point of view this is very positive news as it indicates that the perception of an originally unfamiliar non-native length contrast is indeed learnable (cf. also the supporting results of training studies for Dutch learners of Finnish, e.g., Heeren and Schouten, 2008, for adults and Heeren and Schouten, 2010, for children), at least to a certain extent. Furthermore, increased exposure to the L2, especially also to increased variation in the input, improves listeners' sensitivity (e.g., Sadakata and McQueen, 2011).

Importantly, Morrison (2002) showed that learners may readjust their *perceptual strategies* with increased exposure to the target language: While Spanish beginning learners of English did not show a preferential use of mainly durational or of mainly spectral cues to distinguish between English /i/ vs. /ɪ/, the same learners shifted towards relying on either spectral or durational cues after one month's stay in Canada. This suggests that even in cases where learners' performance (as measured by discrimination or categorization results) approaches that of native speakers of the target language, the perceptual strategies of the two groups may still differ.[3] Such a readjustment of interlanguage phonological systems can not only be observed through improved perceptual performance on L2 contrasts, but also finds expression in longer reaction times compared to target language speakers (e.g., Minagawa-Kawai et al., 2005, for categorical perception of short vs. long vowels, as well as Dijkstra, 2005, for

[3] For an insightful discussion of similar differential auditory cue weighting for /i/ vs /ɪ/ in the interlanguage development of Spanish learners of two varieties of English, see Escudero and Boersma (2004).

lexical decision times). This indicates that the processing cost for L2 learners is higher than for native speakers even when their behavioral data shows (near) target-like performance. The current study aims to investigate these aspects for both consonantal and vocalic length on Italian and German participants within one coherent study, which not only adds a new cross-linguistic dimension but also yields comparable results for consonants and vowels for the same participant groups, which are so far non-existent.

Italian and German length contrasts

Standard German and Italian are both sensitive to quantity, although for different types of segments: German, but not Italian (Krämer, 2009)[4], employs a phonemic vocalic length contrast, e.g. [ban] 'ban' vs. [baːn] 'train', and Italian, but not German, distinguishes between geminate and singleton consonants, e.g. [fato] 'fate' vs. [fatːo] 'fact'. Thus, German and Italian are ideal languages for testing the perceptual sensitivity to length contrasts in a different class of segments, vowels vs. consonants.

The German vowel inventory distinguishes not only between tense and lax vowels, but also between long and short ones (Wiese, 2000).[5] Tense vowels, which

[4] It should be noted that 'Italian' for our purposes is intended to mean Standard Italian. We are aware that certain northern Italian dialects, e.g., Friulian (Baroni and Vanelli, 2000) or Milanese (Prieto, 2000), may contrast short and long vowels.

[5] There are a number of mostly articulatory studies that argue that the tense-lax distinction in German

are produced with more peripheral articulatory settings, on the one hand, appear long in stressed position and short in unstressed position. Lax vowels, the articulatorily more central counterparts, on the other hand, are mostly short. There exist, however, two pairs of lax vowels for which only length is contrastive: /ɛ/ vs. /ɛː/[6] and /a/ vs. /aː/ (Wiese, 2000). For these two vowel pairs length is the only distinguishing feature and independent of stress placement or vowel quality. Ramers' (1988) duration measurements yielded a ratio between short and long vowels from 1.65 to 2.58 (see also Bohn and Flege, 1990). With respect to spectral quality, there is no difference between long and short /a/ in Standard Northern German, while the short front-mid vowel /ɛ/ may be somewhat more centralized compared to its long counterpart (Jørgensen, 1969). Others, however, have shown that this length pair only differs in duration but not in formant frequencies (Bohn and Flege, 1990). Perceptually, German native listeners have been shown to purely rely on durational cues, at least for /a/ vs. /aː/ (Sendelmeier, 1981).[7] Thus, segmental length is clearly important for German native speakers in their vocalic system.

---

reflects a difference in the "Silbenschnitt" and is hence structural rather than purely durational (Hoole and Mooshammer, 2002; Kroos et al., 1997; Vennemann, 1991). Irrespective of the underlying theory, the acoustic difference that the listener has to process for /a:/ vs. /a/ in German manifests itself only in durational terms (Jørgensen, 1969, Pätzold and Simpson, 1997).

[6] We would like to mention that the long front mid-open vowel /ɛː/ is not consistently produced across German dialects. For example, in an empirical study of radio and television news samples, Ramers (1988) found it to be commonly substituted by [eː] by Northern German speakers, while it was present in other varieties. Since our German participants all came from the very south of Germany, they were from a dialect area where there is still an active distinction between [eː] and [ɛː].

[7] See Lehnert-Le Houillier (2007) for a detailed overview of vocalic length studies for German.

The Italian consonantal system contains numerous distinct phonemes that can occur both as singletons and geminates, namely /p/, /b/, /t/, /d/, /k/, /g/, /f/, /v/, /s/, /ʧ/, /ʤ/, /m/, /n/, /l/ und /r/ (Muljačić, 1972). They are contrastive almost exclusively in intervocalic position, but some geminates may also occur before liquids or glides (e.g. [pːr], [pːl], [pːj] or [bːw]). As to the remaining consonants in the inventory, /j, w, z, ʒ/ are always short and /ʃ, ɲ, ʎ, ʦ, ʥ/ are 'intrinsically long' (Passino, 2008).

Acoustic studies investigating the phonetic properties of the singleton/geminate contrast for different consonant types in Italian (stops: Esposito and Di Benedetto, 1999; fricatives: Giovanardi and Di Benedetto, 1998; nasals: Mattei and Di Benedetto, 2000; affricates: Faluschi and Di Benedetto, 2001) revealed that duration is the main property underlying the consonantal contrast. The measurements of intervocalic singleton and geminate consonants from eight Italian native speakers in Kabak et al. (2011) indicate an average duration of 84.5 ms for singletons and of 196 ms for geminates, resulting in a ratio of 2.3. However, the duration of the vowel preceding the geminate or singleton consonant co-varies to some degree: it is longer before a singleton and shorter before a geminate consonant (Esposito and Di Benedetto, 1999). According to Kingston et al. (2009), the allophonic variation in preceding vowel duration is comparatively small (allophonic vowel duration ratio of 1.3-1.5). Perceptually, the geminate-singleton discrimination in Italian depends less on the actual length of the consonant, but rather on the ratio between the consonant and the preceding vowel (Pickett et al. 1999). When the consonant is longer than the preceding vowel, it is perceived as a geminate, if it is shorter as a singleton (Giovanardi and Di Benedetto, 1998), and any following segment has no influence.[8] However, vowel duration is not itself a cue for consonantal length independent of consonant duration (cf. Esposito et al., 1997; Pickett et al., 1999).

In sum, since Italian and German exploit length contrastively in one class of segments but not in the other, this pair of languages is ideally suited to investigate native and non-native length perception in vowels and consonants within one study. Our experiment will provide cross-linguistic data, supplemented by reaction times to compare processing difficulty, which allows us assess underlying processing mechanisms.

Experiment

The research questions under investigation in the current study are the following:

(1) Can listeners with a length contrast in one class of segments (e.g., vowels) employ

[8] It should be added that words with geminate consonants also differ in other dimensions apart from duration. Payne (2006), for instance, reported that length contrasts in coronal sounds (/l/, /n/, /t/, /d/, /s/) also correlate with the exact place of articulation. Specifically, geminate consonants were produced with a more palatalized tongue position, while the tongue was more fronted in singleton consonants.

durational cues for the discrimination of a length contrast in another class of segments (e.g., consonants)?

(2) Does the ability to perceive consonantal length improve with increased exposure to a language with such a contrast? Does the perception of consonantal length contrasts become native-like?

(3) Is it equally difficult to perceive native and non-native vocalic and consonantal length (as estimated by reaction times)?

Given that both kinds of length contrast are signaled by segment durations and given that the non-native durational differences between short and long segments are well beyond the just noticeable difference for duration (and also beyond the differences found for L1 allophonic variation), one might predict that there are no differences in the non-native perception of consonantal and vocalic length contrasts. Bohn (1995) posited the Desensitization Hypothesis, which claims that in the absence of sufficient spectral information listeners will use durational cues to perceive non-native sounds. Although this hypothesis was based on studies on vowel perception, listeners' sensitivity for durational information predicts that both kinds of non-native length contrasts are perceived equally well. According to Escudero's (2005) L2 Linguistic Perception model (L2LP), naïve listeners without experience with a length contrast would be expected to process the respective stimuli initially on a non-linguistic basis, since this model refutes the existence of any native category (i.e., neither a short nor a long category) in the L1 for vowel or consonant duration in such a case. This scenario thus predicts successful discrimination for both types of segments by naïve listeners as well. It must be noted, however, that Escudero's model is dynamic and purely non-linguistic processing only applies to listeners with very little or no L2 experience – in more advanced learners, new categories will have been formed and their boundaries are constantly being adjusted. Thus, after successful (non-linguistic) discrimination at the beginning stage, learners will start processing non-native length contrasts linguistically. This initial stage of linguistic processing may lead to temporarily lower success rates compared to the stage of purely non-linguistic processing. Eventually, the PL2P predicts succinctly improving success rates with increasing L2 exposure, which gradually results in more target-like fine-tuning of the initially quite rough and inaccurate new category boundaries.

Method

Participants. Three groups of participants participated voluntarily in the experiment, 10 speakers of Standard German with no training in Italian (henceforth non-learners), 10 native Germans who were proficient learners of Italian (henceforth German-Italian learners), and 10 native speakers of Italian with no knowledge of

German. None of the participants had any self-reported speech or hearing deficits.

The non-learners consisted of seven female and three male participants, aged between 21 and 28 (mean age: 23.3). None of them had learned Italian or another language with a phonemic consonantal length contrast.

The learners included nine female and one male participant, aged between 20 and 35 (mean age 23.1). They had all studied Italian at the university level for at least 11 months. Except for two, they had all lived in Italy for at least four months after the age of 15. Detailed information regarding the learners' demographics is provided in Table 5 in the Appendix.

The Italian group was comprised of six female and four male participants, aged between 19 and 29 (mean age 24.1). None of them had learned German or stayed in Germany for a longer period, but all except for one had learned English as an L2 (after the age of 10).

Materials. The stimulus material consisted of 41 mono- or disyllabic non-words taken from the Italian GEMMA project (Di Benedetto, 2000), which are phonotactically legal in German and Italian. Four experimental conditions involving segmental and length contrasts, both in vowels and consonants, were created (for an overview see Table 1). For the consonantal length condition, we used VCV and VC:V minimal pairs with an intervocalic contrast between geminates and singletons. For the consonantal segment (control) condition, we used VCV minimal pairs with singleton consonants that only differed in the place of articulation (same manner of articulation and voicing). We used consonants that exist in German and which can occur both as singletons and geminates in Italian. In the consonantal length condition, we included two stimulus pairs with fricatives, two with nasals and two with stop consonants (/f, s, m, n, t, b/). Half of these consonants were phonologically voiced (/m,n,b/), the other half voiceless (/f,s,t/). They were combined with the vowels /a, i, u/, which constitutes a subset of the Italian and German vowel inventories. In the vowel length condition, we used CVC and CVːC minimal pairs with a contrast between long and short vowels. The vowels in the critical vocalic length condition were restricted to /a/, /aː/, /ɛ/ and /ɛː/ since only for these two vowel pairs there is a length contrast in German. In the vowel segment (control) condition, we used CVC minimal pairs with the short vowels /a/ and /ɛ/. The surrounding consonants had the same place and manner of articulation (the voicing of onset and coda consonants sometimes differed due to German syllable-final devoicing of obstruents). Five additional nonwords were created for the practice pairs ([ulu, ulːu, dap, daːp, iri]).

Table 1. Minimal pair contrasts used in the four conditions.

C length	C segment	V length	V segment
[afa] [afːa]	[afa] [asa]	[faf] [faːf]	[faf] [fɛf]
[usu] [usːu]	[usu] [ufu]	[mɛm] [mɛːm]	[mɛm] [mam]
[imi] [imːi]	[imi] [ini]	[bap] [baːp]	[bap] [bɛp]
[unu] [unːu]	[unu] [umu]	[vɛf] [vɛːf]	[vɛf] [vaf]
[ata] [atːa]	[ata] [apa]	[nan] [naːn]	[nɛn] [nan]
[ubu] [ubːu]	[ubu] [ugu]	[gɛk] [gɛːk]	[gɛk] [gak]

Experimental and practice stimuli were read in isolation and recorded twice by a female simultaneous bilingual speaker of German and Italian who had grown up in southern Germany and was exposed to both languages from birth. We used a single speaker for all the stimuli to avoid that listener judgments could be influenced by voice characteristics. Recordings took place in a sound-attenuated cabin at the University of Konstanz. Data were directly digitized using a DAT recording with a sampling rate of 44.1kHz and a resolution of 16Bit.

To verify the durational differences of the stimuli in the length conditions, the respective long and short consonants and vowels were annotated using Praat (Boersma and Weenink, 2009), following standard segmentation criteria (Turk et al., 2006). More specifically, for plosives, we annotated closure duration (starting from a clear drop in the amplitude of the waveform and a drop in higher frequency energy, especially F2, in the spectrogram) up to the release burst and for fricatives the duration of friction (as determined by the presence of aperiodic noise in the waveform). For nasals, measurement started when the amplitude in the waveform dropped and the waveform showed less high frequency components (drop in high frequency energy in spectrogram). Results of an acoustic analysis showed that, on average, long vowels were 2.3 times longer than short vowels ($t(11) = 79.9$, $p < 0.0001$), while geminates were on average 1.9 times longer than singleton consonants ($t(11) = 13.3$, $p < 0.0001$), for details see Table 2). These duration measurements ensured that the acoustic criteria for the length distinction in vowels and consonants were met. Furthermore, the average duration ratios for the critical stimulus segments in the L2 were at least as large as ones reported for allophonic differences in the respective L1s (ratio for German quasi-geminates to singletons: 1.8, cf. Mikuteit (2007), ratio for stressed to unstressed Italian vowels: 2.0, cf. Braun and Geiselmann (2011)).

Table 2. Average durations and standard deviations (in ms) for short and long segments.

Sound	N	Duration and standard deviation (in ms) of short sound	Duration and standard deviation (in ms) of long sound	Duration-Ratio
/a/	6	135.2 (25.4)	294.2 (18.3)	2.2
/ɛ/	6	126.0 (12.1)	298.0 (16.8)	2.4
/b/	2	92.5 (9.2)	177.5 (16.3)	1.9
/f/	2	137.5 (13.4)	232.0 (11.3)	1.7
/m/	2	118.0 (1.4)	225.0 (28.3)	1.9
/n/	2	92.0 (5.7)	191.0 (11.3)	2.1
/s/	2	142.5 (0.7)	255.5 (34.6)	1.8
/t/	2	105.5 (3.5)	184.5 (6.4)	1.7

Furthermore, for vowels, the first and second formants at the mid point of the vowel were automatically extracted. Results of a paired t-test showed a significantly higher F2-value for short /a/ compared to long /aː/ (1435.6 Hz vs. 1322.1 Hz, $t(5) = 4.4$, $p < 0.01$) but no difference in F1 (on average 956 Hz, $p > 0.4$). For /ɛ/, F1 and F2 did not differ for long and short vowels ($p > 0.05$). Average F1 was 624 Hz, average F2 2209 Hz.

To exclude the possibility that hearers could make the critical consonantal length judgment based on the allophonic duration of the vowel preceding the consonant, this vowel was held constant. To this end, for each vowel /a, i, u/ we chose a token that was always shorter in duration than any geminate consonant and longer than any singleton consonant. Using Praat, the selected instances of /a/, /i/ and /u/ were spliced out (at positive zero-crossings) and every vowel preceding a (singleton or geminate) consonant was then replaced by the respective vowel splice. This ensured that the first vowel of every non-word was always constant across all trials in the consonantal length condition.

Procedure. Since every word was recorded twice, there were eight versions of every word pair: For pairs without a contrast (i.e., different tokens of the same type), there were four possible combinations (e.g. afa1-afa2, afa2-afa1, afːa1-afːa2, afːa2-

af:a1)[9], for pairs with a contrast there were eight possible combinations (e.g. afa1-af:a1, af:a1-afa1, afa1-af:a2, af:a2-afa1, afa2-af:a1, af:a1-afa2, afa2-af:a2, af:a2-afa2).

We created two experimental lists, each containing all four versions of the pairs without a contrast and four of the eight versions of the pairs with a contrast. The order of presentation in the pairs with a contrast was counterbalanced across lists (e.g. afa1-af:a1 in list 1 and af:a1-afa1 in list 2). Each list hence contained 192 trials, 24 word pairs x 8 versions (4 with and 4 without a contrast). We created four blocks, containing 48 stimuli each. The experiment started with a training block, consisting of the three practice pairs, two with a length contrast (one vocalic, one consonantal) and one without a contrast. There was no feedback, neither during the practice block nor during the main experimental part. Participants were randomly assigned to one of the lists (five participants per language group for each list).

German participants (non-learners and learners) were tested on the same-different task at the University of Konstanz, either alone or in groups of two to three. The experiment was programmed in C (timing accuracy 1ms) using a microcontroller (80C32, 11.059MHz). Stimuli were presented via headphones (*Sony MDR-CD570*) by means of an external player (*M-audio microtrack II*). Three beeps (1 kHz, 300 ms) signaled the beginning of a block and participants could take breaks between blocks. Each trial started with a 300 ms, 1 kHz sinusoid beep and 500 ms of silence, followed

---

9 The number stands for a particular token of that type, i.e., afa1 and afa2 represent different recordings.

by the first stimulus. The two members of each pair were separated by a silence of 1600 ms. This long inter-stimulus interval was chosen in order to prevent purely acoustic discrimination. Participants were then given a maximum of 1600 ms before timeout Reaction times were measured relative to the offset of the second stimulus and no visual aids were provided. All answers and reaction times were recorded using a button box. Participants used their dominant hand for a 'same' response and their non-dominant hand for a 'different' response.

The Italian participants were tested in Rome. Due to shipping problems with the experimental hardware, the Italian AX-task had to be reprogrammed using Presentation (Neurobehavioral Systems) and the experimental procedure accidentally differed in one aspect from the one for the German participants; Immediately after the second stimulus, Italian participants saw the text *clicca ora* 'click now' on the screen until the end of the trial (1600ms). The timing was otherwise identical and reaction times were in both procedures measured relative to the end of the second stimulus.[10]

We used a speeded AX-discrimination-task in which participants decided as quickly as possible (cf. Babel and Johnson, 2010). AX tasks are ideally suited for non-native listener populations since they do not demand lexical knowledge of the language in question and can hence be performed by learners and non-learners alike. Reaction times have been shown to reveal processing difficulties. For instance, longer

[10] We expect the Italian procedure to result in slightly longer reaction times in all conditions due to the visual prompt. We will address this issue in more detail in the analysis section,

reaction times have been found for pairs that were independently judged to be very similar (Babel and Johnson, 2007). Further support for associating longer reaction times with increased task difficulty is provided by Tomaschek et al. (2011) or Borràs-Comes et al (2010). In these studies, the response time peak coincided with the measured category boundary. Tomaschek et al.'s study is particularly relevant as it showed highest reaction times at the smallest noticeable difference in an experimental investigation of the perception of long vs. short /a/ by German native speakers. Hence reaction times also provide a window into the perceived similarity of stimulus pairs and thus into the relative level of difficulty for participants.

Results

In total, 5760 data points were recorded (30 participants x 192 trials). From these, 75 data points had to be excluded because participants pressed the button before the end of the stimulus (six for non-learners, 69 for learners). Further 218 data points were excluded due to timeout (40 for non-learners, 129 for German-Italian learners, 49 for Italians). Timeout data were analyzed using a binomial logistic regression model with *timeout* (yes/no) as dependent variable and *participant group (*German non-learner, learner or Italian*)*, *type of contrast* (length or segment), *trial type* (without or with contrast) and *segment class* (consonant or vowel) as fixed factors. *Participants* and *items* were added as crossed random factors (Baayen, 2008; Baayen et al., 2008).

Results showed only a main effect of *participant group* ($p < 0.05$): there were more timeouts in the German-Italian L2 group compared to the Italian and German groups.

In the remainder of this section, we will first present the analysis of d' scores, a measure of sensitivity to the kinds of contrasts tested, and then turn to reaction times.

Sensitivity to contrast: d' analysis. The d' scores reflect a participant's sensitivity for differences in the stimuli based on hits and false alarms (Macmillan and Creelman, 2005). We calculated d' scores for each participant for each of the four conditions (consonantal and vocalic length contrasts as well as consonantal and vocalic segment contrasts). Table 3 shows the average d' values for length and segmental contrast in the three participant groups.

Table 3. Mean values and standard deviations for d' scores in the four conditions for each participant group.

	Length contrasts		Segmental contrasts	
	Vowel	Consonant	Vowel	Consonant
German non-learner	2.87	1.31	3.39	3.30
	(0.96)	(0.76)	(0.12)	(0.16)

German Italian L2	2.75	1.95	3.35	3.26
	(0.91)	(0.76)	(0.12)	(0.26)
Italian	3.04	3.04	3.44	3.30
	(0.66)	(0.44)	(0.11)	(0.04)

A general linear model with d' scores as dependent measure and *participant group*, *segment class* (C or V) and *type of contrast* (length or segment) as fixed factors showed a significant three-way interaction ($p < 0.05$).[11] To investigate the nature of this interaction, the data were split by segment class. In the vocalic conditions, there were no effects of *participant group* or *type of contrast* and also no interactions (all p-values > 0.05). The average d' score was 3.14.

In the consonantal conditions, there were main effects of *type of contrast* ($p < 0.0001$) and *participant group* ($p < 0.005$), and an interaction between the two factors ($p < 0.05$), see Figure 1. More specifically, for consonantal length contrasts, d' scores for German non-learners (d' = 1.31) were significantly lower than those for learners (d' = 1.95, $p < 0.05$), which in turn were considerably lower than those for Italians (d' = 3.04, $p < 0.005$). For consonantal segment contrasts, learners (d' = 3.26) did not differ from German non-learners (d' = 3.30, $p > 0.5$); the difference between learners and Italians (d' = 3.44) was significant ($p < 0.05$). Similar to vocalic segment

[11] Due to the limited number of items for each segment in this study, any test for more fine-grained differences in consonants (e.g., plosive, fricatives, nasals) would not be meaningful.

contrasts, all three participant groups' d'-scores were close to ceiling (between 3.26 and 3.44).

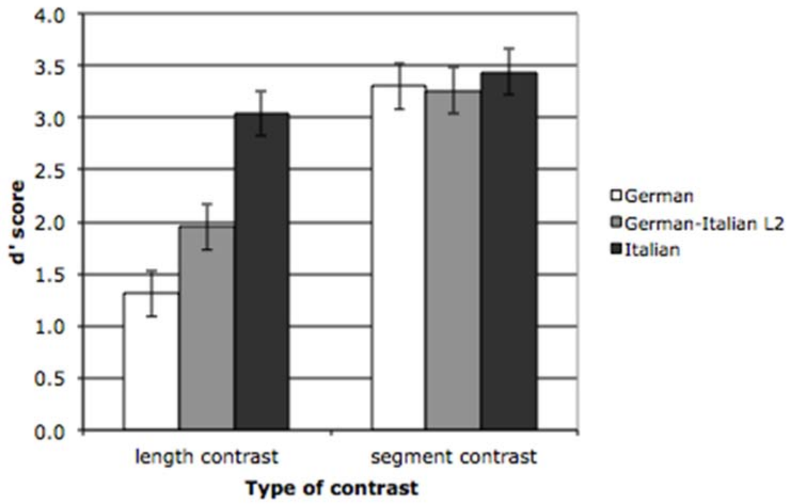

Figure 1. Average d'-scores for the discrimination of consonantal length and segmental contrasts by the three participant groups, based on the statistical model. Whiskers represent standard error.

To test whether the lower d' scores for German non-learners and learners in the consonantal length contrast were caused by difficulties with 'same' or 'different' trials, we also analyzed the accuracy scores in this condition. To this end, we calculated a binomial logistic regression model with *accuracy* as categorical dependent variable and *participant group* and *trial type* ('same' or 'different') as fixed factors (Jaeger, 2008). The results showed main effects of *participant group* ($p < 0.05$) and *trial type*

($p < 0.05$) and a significant interaction between these two factors ($p < 0.05$). To investigate the nature of the interaction between *participant group* and *trial type,* we looked more closely at the reactions to 'same' and 'different' trials, respectively (see Figure 2). For ‘same’ trials, i.e., for trials without a contrast, there were no significant differences across participant groups (all p-values > 0.05). For 'different' trials, however, results showed significant differences across all three groups: Accuracy was higher for German learners of Italian than for German non-learners ($z = 2.0$, $p < 0.05$) and higher for Italians than for German learners ($z = 3.9$, $p < 0.001$). Thus, the difference in d’ scores for consonantal length contrasts between participant groups can be accounted for by their differential performance on ‘different’ trials.

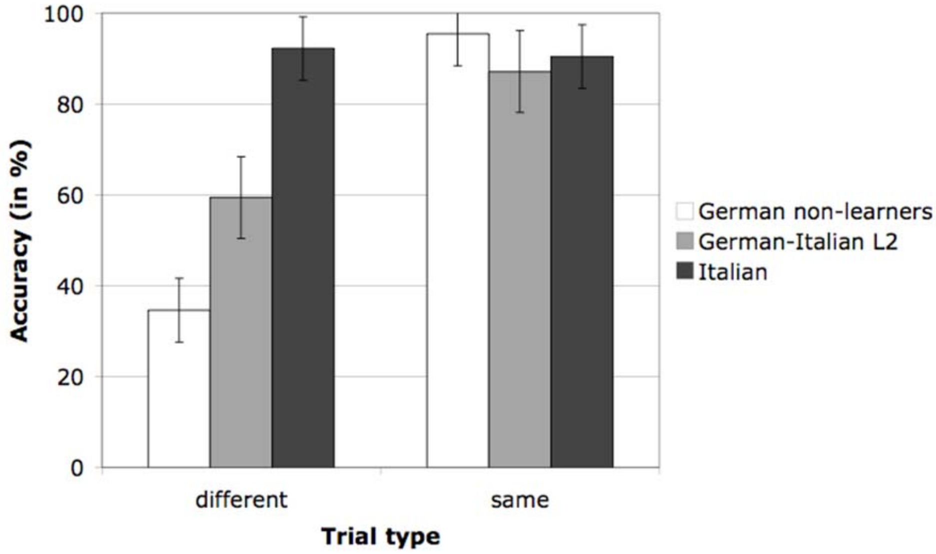

Figure 2. Average accuracy of the three participant groups split by trial type in the consonantal length condition. Whiskers represent standard error.

Processing difficulty: Reaction time analysis. Average reaction times and standard deviations are shown for each condition and each participant group in Table 4.

Table 4. Average reaction times and standard deviations (in ms) for each condition and participant group.

	Length contrasts		Segmental contrasts	
	Vowel	Consonant	Vowel	Consonant
German non-learner	430.1	475.2	407.2	424.4
	(234.0)	(260.9)	(208.6)	(233.9)
German Italian L2	423.3	509.4	385.4	437.8
	(253.5)	(279.4)	(239.4)	(250.1)
Italian	546.6	544.2	458.1	495.9
	(300.6)	(279.4)	(247.6)	(290.1)

As mentioned above, the reaction time (RT) analysis was performed to investigate task difficulty for the different length conditions (native and non-native). We only analyzed RTs in trials with a vocalic and consonantal length contrast[12]. To account for participant-specific RT-differences, we normalized the raw RT data in the following way: RTs for correct responses were averaged for each participant and condition. Then, the (thus averaged) RTs in the segmental control conditions (vowel and consonants) were subtracted from the (averaged) RTs in the respective length

[12] For trials without a length contrast, task difficulty cannot be unambiguously attributed to length.

conditions (vowels and consonants) for each participant.[13] Figure 3 shows the normalized RTs for the three participant groups in the two length conditions. In what follows, we will describe the results of a univariate ANOVA with normalized RT as dependent variable and *participant group* and *segment class* (consonant or vowel) as fixed factors.[14]

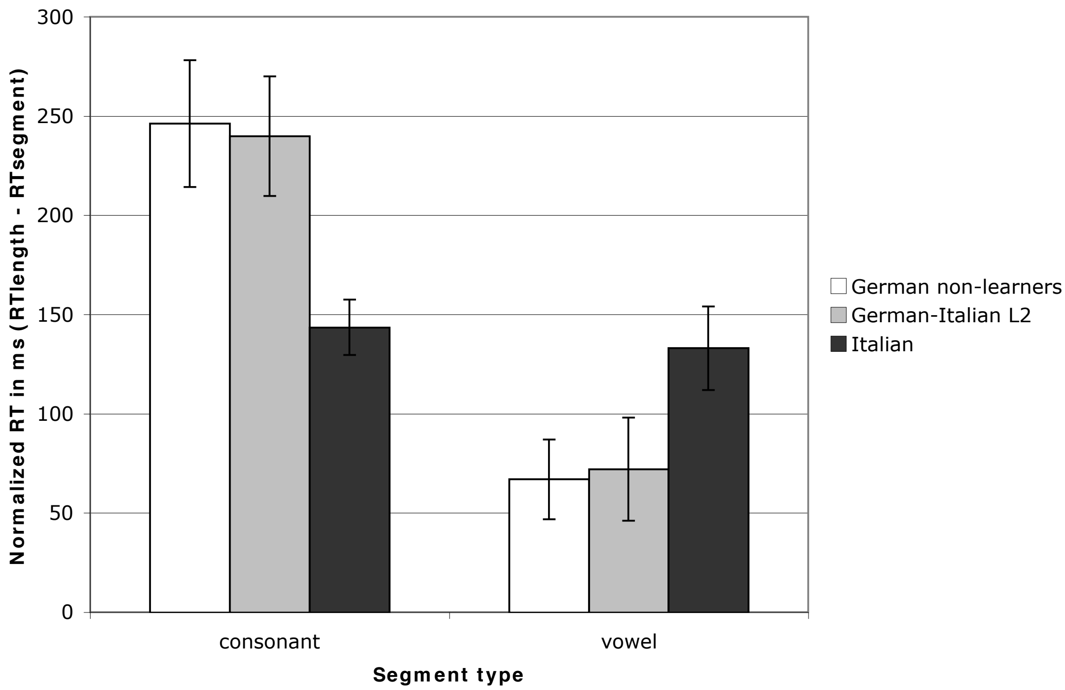

Figure 3. Normalized reaction times (RT length condition - RT segment condition) for

[13] Note that reaction times for segment conditions did not differ as a function of *participant group* or *segment class* (all p-values > 0.2). This suggests that the difference in protocols between German and Italian participants had no effect on reaction times in general.
[14] Since the data were aggregated, we calculated a general linear model instead of a linear mixed effect regression model as done before.

trials with a length contrast as a function of *segment type* and *participant group*. Whiskers represent standard error.

Results showed a main effect of *segment class*: Normalized RTs to trials with a length contrast were significantly shorter for vowel length contrasts than for consonantal length contrasts ($F(1,60) = 34.7$, $p < 0.001$). There was no effect of *participant group* ($p > 0.1$) but an interaction between *participant group* and *segment class* ($F(2,56) = 7.2$, $p < 0.005$). More detailed analyses showed that for Italians there was no difference in RTs for consonants and vowels ($F(1,18) < 1$, $p > 0.8$), while for both German groups reactions were significantly faster for vowel length contrasts than for consonantal length contrasts ($F(1,18) = 22.7$, $p < 0.001$ for non-learners, $F(1,18) = 17.5$, $p = 0.001$ for learners. LSD posthoc tests showed that normalized RTs in the consonantal length condition were significantly longer for the two German groups compared to the Italian group (both p-values < 0.05), while the two German groups did not differ from each other ($p > 0.9$). The difference in normalized RTs in the vowel length condition was also significant ($F(2,27) = 2.8$, $p < 0.05$). German non-learners and learners did not differ from each other ($p > 0.8$), but were significantly faster than Italians ($p < 0.05$, LSD posthoc test).

General Discussion

Analyses of participants' responses in this same-different task clearly showed an asymmetry in the ability to perceive length contrasts that are not phonemic in one's L1. Vowel length contrasts, on the one hand, appeared to pose little problems, even when vowel length was not phonemic in ones' native language: they were perceived as successfully as segmental contrasts. The perception of consonantal length contrasts, on the other hand, proved to be difficult for non-native listeners whose L1 does not distinguish consonantal length. This difficulty still persisted in the advanced group (German learners of Italian). Despite a considerable improvement compared to the group without exposure to a language with consonantal length contrasts, their perception was (still) clearly different from that of native Italian listeners.

We will first focus on the improvement in consonantal length perception, a finding that is in line with recent reports by Heeren and Schouten (2008, 2010). The learners in the current study have had ample experience with Italian (most of them had lived in Italy for at least several months), which is most probably the source for their improved sensitivity to consonantal length contrasts compared to non-learners. In terms of theories of speech perception, there are two possible explanations that can account for the results of the perception of non-native consonantal length contrasts: The learners' phonological system might have restructured in a way to create new categories[15] for each such individual segment that they observe in the target language. More specifically, upon encountering enough instances of a geminate /tː/ or /kː/ as being different from /t/ or /k/, respectively, they establish new categories, and this happens for each single segment.[16] This can be considered a case of input-driven cue-based learning (Dresher and Kaye, 1990; Harrington and Dennis, 2002), which is triggered once a certain frequency threshold of occurrence for each individual consonant has been reached. From an inductive learning perspective, however, it would be more efficient to establish a more general concept of *geminate*, which applies to the whole class of consonants and can thus, once in place, be easily transferred to any new consonant. In order to form a more general concept of *geminate*, learners have to adapt their prelexical processing (McQueen, 2005); only then this (for them) new type of contrast can be processed and integrated into the L2 system. Since the materials were not designed to adjudicate between these two hypotheses, further studies with familiar and novel length contrasts would be needed to resolve this issue.

The findings for improved sensitivity for consonantal length contrasts with

[15] According to Escudero (2005) two new categories for length (short vs. long) have to be created, following Best (1995) it would be one new category.

[16] We assume that this new category formation happens in the sequential order of singleton first - geminate next since short segments are (i) typologically less marked than long ones (Blevins, 20005), i.e. the existence of geminate consonants entails at the very least the existence of their corresponding singletons in a given language inventory and geminates often originated through historical language change (Blevins 2005) from singletons. The actual phonetic duration, of course, may differ between languages.

increased exposure are not easy to explain by current theoretical L2 models as only few go beyond the beginning learners' initial state. Notable exceptions are Flege's (1995) Speech Learning Model SLM (if it were not restricted to segmental features) and Escudero's (2005) L2LP. The SLM implies readjustment of native categories and the formation of new categories in the course of the learning process – in our case the development of a new suprasegmental contrast (length) for a new class of segments (consonants). L2LP also accommodates learning in that newly formed categories are adjusted more and more towards optimal L2 perception.

German speakers with no exposure to a consonantal length contrast obviously do not associate the durational differences in consonants with different phonological categories, as indicated by their low discrimination scores. Thus, they map the consonants – irrespective of their duration – onto one and the same category, the only one they have available in their L1 inventory. This could be interpreted as a native language magnet effect in terms of Kuhl (1995), as single category perceptual assimilation (Best, 1995) or equivalence classification in the SLM. The hypothesized merging of a non-native contrast onto a single category is strengthened by the fact that naïve German listeners mostly misclassified trials with a consonantal length contrast as 'same' and not vice versa (i.e., the obviously different tokens of the same length as 'different'). For this phenomenon, L2LP cannot offer a reasonable explanation to the single category mapping account since this model posits that there is no length category available in the initial stage of this learning process. Consequently, naïve German listeners would be expected to process long and short consonants non-linguistically, which should have led to very good discrimination. Obviously this was not the case for our German non-learners.

Although the German learner group displayed significantly more correct 'different' responses and also a higher sensitivity to the consonantal length contrast (as measured by d') than the naïve German group, the learners' reaction times for such pairs were significantly longer than the reaction times for the native vowel length contrast. What is more, learners' RTs for correctly classified items with a consonantal length contrast were as high as those of non-learners. This suggests that the 'same'-'different' decision (a) was equally difficult for learners as it was for non-learners and (b) has not yet reached the ease with which native vowel length contrasts are processed. Apparently, the learners, who are able to perceive the consonantal length contrast better than the non-learners, do not perceive it reliably as a phonemic contrast, which the long inter-stimulus interval in the current study required. Despite their extended exposure to consonantal length contrasts in the target language, they still display some level of indecisiveness as illustrated by long RTs in the consonantal length condition and also by the high number of timeouts (more than in both other groups). This indecisiveness suggests that (at least some of) the learners already had some access to a newly forming category and thus needed to decide between two possible mapping representations (i.e., /C/ vs. /Cː/). The fact that the naïve German group also displayed quite high RTs (but fewer timeouts) may be explained by a difference in 'category goodness' (in terms of Best's (1995) Perceptual Assimiliation Model) of one of the consonants in question, which also makes discrimination more difficult. However, the underlying mechanisms for non-learners and learners need further attention in future experiments.

As for the perception of the vocalic length contrast, all participant groups achieved a high level of accuracy in the current study, regardless of the native language. As illustrated by the d' and accuracy scores, the non-native length contrast in vowels did not pose any problems for the Italian listeners whose native language system does not include a vowel length contrast. These findings corroborate previous accounts in the literature (e.g., Bohn, 1995; Cebrian, 2006; Flege et al., 1997; García Lecumberri and Cenoz, 1997) that the existence of L1 vowel length contrasts is not a necessary condition for the perception of L2 vowel length contrast. Nevertheless it is striking that Italians with no exposure to German behaved native-like in the discrimination of vocalic length. We see a number of possible explanations for the higher sensitivity to vocalic length, which are not necessarily related to the L1 phonological system though. First, the long /a/ was always more fronted than the short /a/, as evidenced by different F2 values in the acoustic analysis. Therefore, Italian listeners might have theoretically relied on spectral information instead of vowel duration. This explanation is rather weak for several reasons: (a) Previous studies have shown that L2 learners weigh duration more than spectral information (Bohn 1995) and that naïve listeners are highly successful for non-native durational contrasts along that dimension (Escudero and Boersma, 2004). (b) The spectral differences are minimal and probably mapped onto the available Italian segmental L1 category (disregarding durational differences) and (c) the spectral difference was only present in one of the two vowels tested. Synthetic stimuli would be needed, however, to entirely refute this confounding factor. Second, the Italian participants all had some learning experience with English, a language that also involves long and short vowels (in combination with spectral differences). Unfortunately, it is impossible to find a group of Italians that is (a) comparable in age and education to the German learners of Italian and (b) has no experience with English. Italian late learners' pronunciation problems with such English vowel contrasts (e.g., *pitch* vs. *peach*) are, however, well documented in the literature, both temporally and spectrally (e.g., Munro et al., 1996; Flege et al., 1999) and successful acquisition of this vowel contrast is mostly limited to Italian speakers who had arrived at an early age (< 10 years) in an English speaking environment (e.g., Munro et al., 1996; Flege et al., 1999). Since the Italian participants in the current study all lived in Italy and started learning English in a formal foreign language setting, i.e., in the Italian school system, after the age of 10 a strong positive influence of English is doubtful. Third, one might argue that the sensitivity of the Italian group to the non-native vowel length contrasts could also be attributed to their experience with allophonic vowel duration: In Italian, stressed open syllables especially in penultimate position are considerably lengthened (Bertinetto, 1980; Braun and Geiselmann, 2011; Krämer, 2009). If experience with allophonic vowel duration were the main predictor for the ability to perceive length contrasts, however, then German listeners would be expected to perform equally well as Italian listeners since they also have experience with allophonic consonant length. As already mentioned in the Background section, the ratio between German quasi-geminates and singletons is comparable to the geminate-singleton ratio in the current stimuli. On the other hand, allophonic consonant duration in German occurs much less frequently than allophonic vowel duration in Italian and, as Kondaurova and Francis (2008) have shown, the frequency of occurrence of allophonic durational differences is a predictor for how strongly listeners rely on duration in a non-native categorization task. In future studies it will be interesting to investigate more closely under which conditions allophonic durational differences are processed efficiently in the native language (for example, if there is a frequency threshold) and whether this reliably predicts success in length distinctions in a second language. Fourth, the observed asymmetry in sensitivity to length contrasts could be also due to the different duration ratios for consonantal and vocalic length contrasts. As shown in Table 2, vocalic length contrasts had a larger duration ratio than consonantal length contrasts (2.3 vowel and

1.9 for consonants). However, both consonantal and vocalic length contrasts lie within the range of L1 allophonic duration variation observed in the two classes of sounds (vowels and consonants) and therefore should not advantage one particular language group. Finally, vowels carry a high informational load. For instance, rhythm and intonation are mainly implemented by suprasegmental features (duration, amplitude, f0) realized on vowels. Not the least to mention in this context is word stress, which is contrastive in Italian and for which duration is the most reliable cue (e.g. Bertinetto, 1980). Furthermore, vowels also contain cues to neighboring segments (e.g., formant transitions) and may convey speaker identity. Therefore, vowels might just be processed more efficiently than consonants in general (cf. Cutler and Mehler's (1993) 'periodicity bias' for infants). Weighting these alternative interpretations, we consider the nature of the acoustic signal and its functional load for different lexical and pragmatic purposes to be one of the most convincing reasons for the successful discrimination of vocalic length in the Italian group.

An additional argument for an asymmetric processing rests on the observation that novel consonantal length contrasts are not only less successful but also more difficult to perceive in an L2 compared to novel vocalic length contrasts. To be precise, the reaction times for Italians in the vowel length condition were not longer than their reaction times on the native consonantal length contrast, which suggests that both contrasts were equally easy or difficult to respond to. On the other hand, for all

German participants (learners and non-learners alike), the non-native consonantal length condition induced longer reaction times compared to the native vocalic length condition. This asymmetry in reaction times is a strong indicator that the two types of non-native length contrast were not equally easy to process.

We thus see a learning development that manifests itself only in terms of accuracy but not (yet) in terms of processing speed. This points to the following asymmetric scenario: (1) Novel length contrasts for consonants are difficult to process but can be improved with increasing L2 experience. Nevertheless, they remain more difficult to process than native vowel length contrasts. (2) Novel vowel length contrasts do not pose a problem even in the absence of experience – neither regarding accuracy nor processing speed.

Conclusion

Listeners' durational sensitivity to non-native vowel length contrasts seems to be high even when vowel length is not contrastive in one's native language. Further psychophonetic experiments are necessary to delimit the bounds of this perceptual ability. On the contrary, consonantal length contrasts remain to be a challenge even after prolonged exposure to (and increased proficiency in) a language that employs such contrasts. It remains unclear how important (the frequency of occurrence of) allophonic experience with durational contrasts for a given class of segments – either consonants or vowels – in the L1 may be. Conceivably, the observed cross-linguistic asymmetry in the ability to perceive non-native vowel and consonantal length contrasts may have its roots in the different informational value carried by vowels and consonants, respectively.

References

Abel SM (1972) Duration discrimination of noise and tone bursts. *Journal of the Acoustical Society of America* 51: 1219-1223.

Altmann H (2006) *The perception and production of second language stress: A crosslinguistic eperimental study*. Doctoral Dissertation. University of Delaware.

Altmann H and Kabak B (2011) Second language phonology. In: Botma B, Kula NC and Nasukawa K (eds) *The Continuum Companion to Phonology*. London: Continuum, pp. 298-319.

Baayen HR (2008) *Analyzing linguistic data. A practical introduction to statistics using R*. Cambridge: Cambridge University Press.

Baayen HR, Davidson DJ and Bates DM (2008) Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language* 59(4): 390-412.

Babel M and Johnson K (2010) Accessing psycho-acoustic perception and language-specific perception with speech sounds. *Laboratory Phonology* 1; 179-205.

Baroni M and Vanelli L (2000) The relationship between vowel length and consonantal voicing in Friulian. In: Repetti L (ed) *Phonological theory and the dialects of Italy*. Amsterdam: John Benjamins, pp. 13-44.

Baumann S, Grice M and Steindamm S (2006) Prosodic marking of focus domains - categorical or gradient? Paper presented at the 3rd International Conference on Speech Prosody, Dresden, Germany.

Bertinetto PM (1980) The perception of stress by Italian speakers, *Journal of Phonetics* 8: 385-395.

Best CT (1995) A direct realist perspective on cross-language speech perception. In: Strange W (ed) *Speech perception and linguistic experience: Issues in cross-language speech research*. Timonium, MD: York Press, pp. 167–200.

Blevins J (2005) The typology of geminate inventories: Historical explanations for recurrent sound patterns. *Seoul Linguistics Forum 2005*. Seoul: Language Education Institute, Seoul National University, pp. 121-137.

Boersma P and Weenink D (2009) *Praat: Doing phonetics by computer*. [Computer program]. Available at: http://www.praat.org/

Bohn OS (1995) Cross-language speech perception in adults: first language transfer doesn't tell it all. In: Strange W (ed) *Speech perception and linguistic experience: Issues in cross-language research*. Timonium, MD: York Press, pp. 279-304.

Bohn OS and Flege JE (1990) Interlingual identification and the role of foreign language experience in L2 vowel perception. *Applied Psycholinguistics* 11: 303-328.

Borràs-Comes J, Vanrell MM and Prieto P (2010) The role of pitch range in establishing intonational contrasts in Catalan. *Proceedings of the 5th International Conference on Speech Prosody*. Chicago: 1-4.

Braun B and Geiselman S (2011) Italian in the no man's land between stress-timing and syllable-timing? Speakers are more stress-timed than listeners. *Proceedings of 12th Annual Conference of the International Speech Communication Association (Interspeech)*. Florence, Italy: 2697-2700.

Braun B, Dainora A and Ernestus M (2011a) An unfamiliar intonation contour slows down on-line speech comprehension. *Language and Cognitive Processes* 26(3): 350-375.

Braun B, Lemhöfer K and Mani N (2011b) Perceiving unstressed vowels in foreign-accented English. *Journal of the Acoustical Society of America* 129: 376-387.

Brown C (1998) The role of the L1 grammar in the L2 acquisition of segmental structure. *Second Language Research* 14: 136-193.

Brown C (2000) The interrelation between speech perception and phonological acquisition from infant to adult. In: Archibald J (ed) *Second Language*

Acquisition and Linguistic Theory. Oxford: Blackwell, pp. 4-63.

Cebrian J (2006) Experience and the use of duration in the categorization of L2 vowels. *Journal of Phonetics* 34: 372-387.

Chiao WH, Kabak B and Braun B (2011) When more is less: Non-native perception of level tone contrasts. *Psycholinguistic Representation of Tone Conference*, Hong Kong, China: 42-45.

Cho T and McQueen J (2005) Prosodic influences on consonant production in Dutch: Effects of prosodic boundaries, phrasal accent and lexical stress. *Journal of Phonetics* 33(2): 121-157.

Creelman CD (1962) Human discrimination of auditory duration. *Journal of Acoustical Society of America* 34: 582-593.

Cutler A and Mehler J (1993) The periodicity bias. *Journal of Phonetics* 21: 103-108.

Delattre P (1966) A comparison of syllable length conditioning among languages. *International Review of Applied Linguistics* IV: 193-198.

Di Benedetto MG (2000) Gemination in Italian: the GEMMA project. *Journal of the Acoustical Society of America* 108: 2507.

Dijksta T (2005) Bilingual visual word recognition and lexical access. In: Kroll FF and De Groot AMB (eds) *Handbook of bilingualism: Psycholinguistic approaches*. Oxford: Oxford University Press, pp. 179-201.

Dresher BE and Kaye JD (1990) A computational learning model for metrical phonology. *Cognition* 34: 137-195.

Dunn I (1993) *The phonetics and phonology of geminate consonants: a production study*. Doctoral Dissertation. Yale University, New Haven.

Dupoux E, Sebastián-Gallés N, Navarrete E and Peperkamp S (2008) Persistent stress 'deafness': The case of French learners of Spanish. *Cognition* 106(2): 682-706.

Escudero P (2005) *Linguistic Perception and Second Language Acquisition*. Doctoral Dissertation. Utrecht University.

Escudero P and Boersma P (2004) Bridging the gap between L2 speech perception research and phonological theory. *Studies in Second Language Acquisition* 26: 551-585.

Escudero P, Hayes-Harb R and Mitterer H (2008) Novel second-language words and asymmetric lexical access. *Journal of Phonetics* 36: 345-360.

Esposito A and Di Benedetto MG (1999) Acoustic and perceptual study of gemination in Italian stops. *Journal of the Acoustical Society of America* 106: 2051-2062.

Faluschi S and Di Benedetto MG (2001) Acoustic analysis of singleton and geminate affricates in Italian. *WEB-SLS, The European Student Journal of Language and Speech*. Available at: http://www.essex.ac.uk/web-sls/papers/01-01/submission.htm

Esposito A, Di Benedetto MG and Rossetti R (1997) Acoustical and perceptual study of gemination in Italian stop consonants. Unpublished manuscript.

Flege JE (1995) Second language speech learning: Theory, findings and problems. In: Strange W (ed) *Speech perception and linguistic experience: Issues in cross-language speech research*. Timonium, MD: York Press, pp. 233–272.

Flege JE, Bohn OS and Jang S (1997) Effects of experience on non-native speakers' production and perception of English vowels. *Journal of Phonetics* 25: 437-470.

Flege JE, MacKay IRA and Meador D (1999) Native Italian speakers' perception and production of English vowels. *Journal of the Acoustical Society of America* 106: 2973-2987.

Fougeron C and Keating PA (1997) Articulatory strengthening at edges of prosodic domains. *Journal of the Acoustical Society of America* 101(6): 3728-3740.

Gandour J (1983). Tone perception in Far Eastern languages. *Journal of Phonetics* 11: 149-175.

García Lecumberri ML and Cenoz Inagui J (1997) L2 perception of English vowels: Testing the validity of Kuhl's prototypes. *Revista Alicantina de Estudios Ingleses* 10: 55-68.

Giovanardi M and Di Benedetto MG (1998) Acoustic analysis of singleton and geminate fricatives in Italian. *WEB-SLS, The European Student Journal of Language and Speech*. Available at:

http://www.essex.ac.uk/web-sls/papers/98-01/98-01.html

Gottfried T and Beddor P (1988) Perception of temporal and spectral information in French vowels. *Language and Speech* 31: 57–75.

Han M (1992) The timing control of geminate and single stop consonants in Japanese: A challenge for nonnative speakers. *Phonetica* 49: 102-127.

Hardison DM and Saigo MM (2010) Development of perception of second language Japanese geminates: Role of duration, sonority, and segmentation strategy. *Applied Psycholinguistics* 31: 81-99.

Harrington M and Dennis S (2002) Input-driven language learning. *Studies in Second Language Acquisition* 24: 261-268.

Hayes R (2002) The perception of novel phoneme contrasts in a second language: A developmental study of native speakers of English learning Japanese singleton and geminate consonant contrasts. *University of Arizona Coyote Working Papers, 12: Language in Cognitive Science*. Tucson: University of Arizona, pp. 28–41.

Hayes-Harb R (2005) Optimal L2 speech perception: native speakers of English and Japanese consonant length contrasts. *Journal of Language and Linguistics* 4: 1–29.

Hayes-Harb R and Masuda K (2008) Development of the ability to lexically encode novel second language phonemic contrasts. *Second Language Research* 24: 5-

33.

Heeren WFL and Schouten MEH (2008) Perceptual development of phoneme contrasts: How sensitivity changes along acoustic dimensions that contrast phoneme categories. *Journal of the Acoustical Society of America* 124: 2291–2302.

Heeren WFL and Schouten MEH (2010) Perceptual development of the Finnish /t-t:/ distinction in Dutch 12-year-old children: A training study. *Journal of Phonetics* 38: 594-603.

Hisagi M, Shafer VL, Strange W and Sussman ES (2010) Perception of a Japanese vowel length contrast by Japanese and American English listeners: behavioral and electrophysiological measures. *Brain Research* 1360: 89-105.

Hoole P and Mooshammer C (2002) Articulatory analysis of the German vowel system. In: Auer P, Gilles P and Spiekermann H (eds), *Silbenschnitt und Tonakzente*. Tübingen: Niemeyer, pp. 129-152.

House AS and Fairbanks G (1953) The influence of consonants' environment upon the secondary acoustical characteristics of vowels. *Journal of the Acoustical Society of America* 25: 1.

Ioup G (1984) Is there a structural foreign accent? A comparison of syntactic and phonological errors in second language acquisition. *Language Learning* 34(2): 1-15.

Jacquemot C, Pallier P, LeBihan D, Dehaene S and Dupoux E (2003) Phonological grammar shapes the auditory cortex: A functional magnetic resonance imaging study. *Journal of Neuroscience* 23: 9541-9546.

Jaeger TF (2008) Categorical data analysis: Away from ANOVAs (transformation or not) and towards logit mixed models. *Journal of Memory and Language, 59*(4): 434-446.

Jørgensen H (1969) Die gespannten und ungespannten Vokale in der norddeutschen Hochsprache mit einer spezifischen Untersuchung der Struktur der Formantenfrequenzen [Tense and lax vowels in Northern German with a focus on formant frequencies]. *Phonetica* 19: 217-245.

Kabak B, Reckziegel T and Braun B (2011). Timing of second language geminates and singletons. *Proceedings of the 17th International Congress of the Phonetic Sciences*: 994-997.

Kijak A (2009) *How stressful is L2 stress? A cross-linguistic study of L2 perception and production of metrical systems*. Doctoral Dissertation. Utrecht University.

Kingston J, Kawahara S, Chambless D, Mash D, and Brenner-Alsop E (2009) Contextual effects on the perception of duration. *Journal of Phonetics* 37: 297-320.

Kondaurova MV and Francis AL (2008) The relationship between native allophonic experience with vowel duration and perception of the English tense-lax vowel contrast by Spanish and Russian listeners. *Journal of the Acoustical Society of America* 124(6): 3959–3971.

Krashen S (1985) *Language Acquisition and Language Education*. San Francisco: Alemany Press.

Krämer M (2009) *The phonology of Italian*. Oxford: Oxford University Press.

Kroos C, Hoole P, Kühnert B and Tillmann HG (1997) Phonetic evidence for the phonological status of the tense-lax distinction in German. *Forschungberichte des Instituts für Phonetik und Sprachliche Kommunikation der Universität München (IPKM)* 35: 17-25.

Kuhl PK (1993) Innate predispositions and the effects of experience in speech perception: The native language magnet theory. In: De Boysson-Bardies B. (ed) *Developmental Neurocognition: speech and face processing in the first year of life*. The Hague: Kluwer Academic Publishers, pp. 259-74.

Ladefoged P and Maddieson I (1996) *The sounds of the world's languages*. Oxford: Blackwell.

Lehiste I (1972) The timing of utterances and linguistic boundaries. *Journal of the Acoustical Society of America* 51: 2018–2024.

Lehnert-Le Houillier H (2007) *The perception of vowel quantity: a cross-linguistic investigation*. Doctoral Dissertation. State University of New York at Buffalo.

McAllister R, Flege JE and Piske T (2002) The influence of L1 on the acquisition of

Swedish quantity by native speakers of Spanish, English and Estonian. *Journal of Phonetics* 30: 229-258.

McQueen JM (2005) Speech perception. In: Lamberts K and Goldstone R (eds) *The handbook of cognition*. London: Sage Publications, pp. 255-275.

Macmillan NA and Creelman CD (2005) *Detection Theory. A user's guide*. Mahwah, NJ: Lawrence Erlbaum.

Mah J and Archibald J (2003) Acquisition of L2 length contrasts. In: Liceras JM, Zobl H and Goodluck H. (eds) *Generative Approaches to Second Language Acquisition (GASLA 2002)*. Somerville, MA: Cascadilla Press, pp. 208-212.

Major, RC (2002) The ontogeny and phylogeny of second language phonology. In: Leather J and James A (eds) *New Sounds 2000: Proceedings of the 4th International Symposium on the Acquisition of Second-Language Speech*. Klagenfurt: University of Klagenfurt, pp. 223–235.

Mattei M and Di Benedetto MG (2000). Acoustic analysis of singleton and geminate nasals in Italian. *WEB-SLS, The European Student Journal of Language and Speech*. Available at: http://www.essex.ac.uk/web-sls/papers/00-02/00-02.html

Mikuteit S (2007) *A cross-language approach to voice, quantity and aspiration: An East-Bengali and German production Study*. Doctoral Dissertation. University of Konstanz.

Minagawa-Kawai Y, Mori K and Sato Y (2005) Different brain strategies underlie the categorical perception of foreign and native phonemes. *Journal of Cognitive Neuroscience* 17(9): 1376-1385.

Morrison GS (2002) Perception of English /i/ and /□/ by Japanese and Spanish Listeners: Longitudinal results. *Proceedings of the North West Linguistics Conference 2002*: 29–48.

Muljaćič Ž (1972) *Fonologia della Lingua Italiana [Italian phonology]*. Bologna: Il Mulino.

Munro MJ and Derwing TM (1995) Foreign accent, comprehensibility, and intelligibility in the speech of second language learners. *Language Learning* 45(1): 73-97.

Munro MJ, Flege JE and MacKay IRA (1996) The effects of age of second-language learning on the production of English vowels. *Applied Psycholinguistics* 17: 313-334.

Nenonen S, Shestakova A, Houtilainen M. and Näätänen R. (2003) Linguistic relevance of duration within native language determines the accuracy of speech-sound duration processing. *Cognitive Brain Research* 16: 492–95.

Nenonen S, Shestakova A, Houtilainen M and Näätänen R (2005) Speech-sound duration processing in a second language is specific to phonetic categories. *Brain and Language* 92(1): 26-32.

Nimz K (2011) Vowel perception and production of late Turkish learners of L2

German. *Proceedings of the 17th International Congress of the Phonetic Sciences*: 1494-1497.

Passino D (2008) *Aspects of consonantal lengthening in Italian*. Padova: Unipress.

Pätzold M and Simpson A (1997) Acoustic analysis of German vowels in the *Kiel Corpus of Read Speech*. In: Simpson AP, Kohler KJ, Rettstadt T (eds) *The Kiel Corpus of Read/Spontaneous Speech - Acoustic data base, processing tools and analysis results*, Arbeitsberichte des Instituts für Phonetik und digitale Sprachverarbeitung der Universität Kiel (AIPUK) 32, pp. 215-247.

Payne E (2006) Non-durational indices in Italian geminate consonants. *Journal of the International Phonetic Association* 36: 83–95.

Peterson GE and Lehiste I (1960) Duration of syllable nuclei in English. *Journal of the Acoustical Society of America* 32: 693-703.

Pickett ER, Blumstein SE and Burton MW (1999) Effects of speaking rate on the singleton/geminate consonant contrast in Italian. *Phonetica* 56: 135-157.

Pohl M and Grijzenhout J (2010) Phrase-medial bilabial stops in three West Germanic languages. *Linguistische Berichte* 222: 141-167.

Prieto P (2000) Vowel lengthening in Milanese: A case for prosodic optimization. In: Repetti L (ed) *Phonological theory and the dialects of Italy*. Amsterdam: John Benjamins, pp. 255-272.

Port RF (1981) Linguistic timing factors in combination. *Journal of the Acoustical*

Society of America 69: 262-274.

Ramers KH (1988) *Vokalquantität und –qualität im Deutschen [Vowel quantity and quality in German]*. Tübingen: Max Niemeyer.

Sadakata M and McQueen JM (2011). The role of variability in non-native perceptual learning of a Japanese geminate-singleton fricative contrast. *Proceedings of the 12th Annual Conference of the International Speech Communication Association (Interspeech):* 873-876.

Salverda AP, Dahan D and McQueen JM (2003) The role of prosodic boundaries in the resolution of lexical embedding in speech comprehension. *Cognition* 90: 51-89.

Sendelmeier W (1981) Der Einfluß von Qualität und Quantität auf die Perzeption betonter Vokale des Deutschen [The influence of quality and quantity on the perception of stressed vowels in German]. *Phonetica* 38: 291-308.

Tagliapietra L and McQueen JM (2010) What and where in speech recognition: Geminates and singletons in spoken Italian. *Journal of Memory and Language* 63: 306-323.

Tomaschek F, Truckenbrodt H and Hertrich I (2011) Processing German vowel quantity: Categorial perception or perceptual magnet effect? *Proceedings of the 17th International Congress of the Phonetic Sciences*: 2002-2005.

Tremblay A (2008) Is second language lexical access prosodically constrained?

Processing of word stress by French Canadian second language learners of English. *Applied Psycholinguistics* 29: 553-584.

Turk A, Nakai S and Sugahara M (2006) Acoustic segment durations in prosodic research: A practical guide. In: Sudhoff S, Lenertová D, Meyer R, Pappert S, Augurzky P, Mleinek I, Richter N, & Schließer J (eds), *Methods in empirical prosody research*. Berlin: Walter de Gruyter, pp. 1-27.

Turk A and Shattuck-Hufnagel S (2007) Multiple targets of phrase-final lengthening in American English. *Journal of Phonetics* 35: 445-472.

Vennemann T (1991) Syllable structure and syllable cut prosodies in Modern Standard German. In: Bertinetto PM (ed) *Certamen Phonologicum II: Papers from the 1990 Cortona Phonology Meeting*. Torino: Rosenberg & Sellier, pp. 211-243.

Wang X and Munro MJ (1999) The perception of English tense-lax vowel pairs by native Mandarin speakers: The effect of training on attention to temporal and spectral cues. *Proceedings of the 14th International Congress of Phonetic Sciences*: 125-129.

Wang Y, Jongman A and Sereno JA (2003) Acoustic and perceptual evaluation of Mandarin tone productions before and after perceptual training. *Journal of the Acoustical Society of America* 113: 1033-1044.

Weber A and Cutler A (2004) Lexical competition in non-native spoken-word recognition. *Journal of Memory and Language* 50(1): 1-25.

White L (1989) *Universal grammar and second language acquisition*. Amsterdam: John Benjamins.

White L (2003) *Second language acquisition and Universal Grammar*. Cambridge: University Press.

White L and Turk A (2010) English words on the Procrustean bed: Polysyllabic shortening reconsidered. *Journal of Phonetics* 38: 459–471.

Wiese R (2000) *The phonology of German*. Oxford: Oxford University Press.

Wightman CW, Shattuck-Hufnagel S, Ostendorf M and Price PJ (1992) Segmental durations in the vicinity of prosodic phrase boundaries. *Journal of the Acoustical Society of America* 91: 1707-1717.

Ylinen S, Shestakova A, Alku P and Huotilainen M (2005) The perception of phonological quantity based on durational cues by native speakers, second-language users and nonspeakers of Finnish. *Language and Speech* 48: 313-338.

Appendix

Table 5. Information on the German L2 learners of Italian

Participant	Age	Gender	Experience with Italian	Age of first contact	Length of Stay in Italy
l1	24	m	3 years at school, 1 year at university	15	5 ½ months
l2	23	w	1½ years at university	20	1 year
l3	22	w	3 years at university	19	none
l4	22	w	4 years at school, 3 years at university	15	10 months
l5	35	w	3 years at Italian university (at the age of 25)	25	3 years
l6	21	w	7 months at school in Italy, 4 months at university	16	7 months
l7	20	w	3 years at school (at age of 15), 4 months at university	15	none
l8	22	w	5 years at school (at age of 15), 2 years at university	15	4 months
l9	29	w	11 years at	18	2 years

			university, (married to an Italian)		
l10	25	w	At school and at university (Italian boyfriend)	18	1 year