Features of low functional load in mono- and bilinguals' lexical access: Evidence from Swedish tonal accent

Nadja Althaus[1,2,*], Allison Wetterlin[2,3] & Aditi Lahiri[2]

[1] School of Psychology, University of East Anglia, Norwich, UK

[2]Faculty of Linguistics, Philology and Phonetics, University of Oxford, Oxford, UK

[3]Department of Foreign Languages and Translation, University of Agder, Kristiansand, Norway

AUTHOR ACCEPTED MANUSCRIPT

Short title:
Swedish tonal accent in mono- and bilinguals' lexical access

*Corresponding author:
Nadja Althaus
School of Psychology
University of East Anglia
Norwich Research Park
Norwich, NR4 7JT
United Kingdom
n.althaus@uea.ac.uk

Abstract

Swedish makes use of tonal accents (Accents 1 and 2) to contrast words, but the functional load is very low, with some regional dialects not even exhibiting the contrast. In particular given the low number of minimal pairs, the question is whether tonal word accent is used in lexical access. Here we present two cross-modal fragment semantic priming studies in order to address this question. Both experiments use first syllable fragments in order to prime semantically related targets. Experiment 1 utilises words whose first syllable occurs with both accent patterns, creating a situation in which there is lexical competition between words that differ solely in terms of accent. Experiment 2 removes this competition by using words that have no such accent competitors. Our results show that native speakers of Swedish use tonal word accent in lexical access: Accent mispronunciations failed to prime semantically related targets, regardless of whether primes had accent competitors or not. Results for a group of early bilingual speakers (who grew up with one Swedish-speaking parent and one other non-tonal language) showed no differences in processing compared to the monolinguals. This indicates that the extraction of accent features during acquisition and their use in lexical access is robust, even in a scenario where multiple input languages lead to tonal word accent being a useful feature for only some of the lexical items that are being acquired. There is no doubt that the accent system is well entrenched into the bilinguals' phonological system.

Keywords: bilingualism; lexical access; lexical tone; low functional load; pitch accent; semantic priming

Features of low functional load in mono- and bilinguals' lexical access: Evidence from Swedish tonal accent

Some linguistic features exhibit low functional load in the sense that they distinguish only few minimal pairs. Do listeners use such features in lexical access, despite their apparent redundancy? Here, we investigate the case of tonal word accent in Swedish. Like Norwegian, Swedish exhibits two accent patterns (Accent 1 and 2). In many tonal languages such as the Asian (e.g., Sino-Tibetan) and African (e.g. Bantu) languages, tonal accent carries high functional load (Gussenhoven, 2004). Segmental information by itself is not sufficient for lexical identity, and previous research has shown that for (monolingual) native speakers lexical access is constrained by suprasegmental information (Fox & Unkefer, 1985; Cutler & Chen, 1997; see below). In Swedish, however, the functional load of tonal word accent is considered to be low. There are only a few hundred minimal pairs (Elert, 1972, 1981), e.g. *anden*$_1$ 'the duck' – *anden*$_2$ 'the ghost, spirit'; *tomten*$_1$ 'the plot, land' – *tomten*$_2$ 'the gnome/Father Christmas'. As these examples demonstrate, minimal pairs often only contrast when the definite article *-(e)n/-(e)t* is added. Moreover, some regional variants do not use accent contrasts at all (e.g. Swedish spoken in Finland). Nevertheless, Stockholm Swedish native speakers use the accent contrast without fail, and furthermore readily distinguish between words on the basis of accent (Felder, Jönsson-Steiner, Eulitz, & Lahiri, 2009). The question is therefore whether tonal accent is, despite the low functional load, such an intrinsic aspect of lexical processing that it is useful in lexical access, or whether speakers instead dismiss this potential cue, relying entirely on segmental information, which almost always is sufficient for word recognition.

Here, we use a cross-modal fragment semantic priming task with accent mispronunciations to examine whether tonal accent plays a role in Swedish native speakers' lexical access. In particular, we ask here whether tonal accent restricts lexical activation. Our paradigm involves priming a target for a visual lexical decision task with an auditory first-syllable prime fragment. For example, the target *häst* 'horse' might be preceded by the fragment [pɔn$_1$], which is the first syllable of the semantically related *ponny*$_1$ 'pony'. If tonal accent plays a role in lexical access, then there should be a difference between accurately pronounced first syllable fragments and versions that are identical in terms of segments but mispronounced with regard to accent. In particular, priming should be restricted to accurately pronounced first syllable fragments, and should not occur for mispronunciations.

One factor that might be important in this context, however, is lexical competition. In Experiment 1 we use cases where there is competition, i.e. where the first syllable could stem from either one of two words as far as segments are concerned, but tonal accent will allow discrimination. We contrast this with cases where accent information is truly redundant because there are no other words in the lexicon that begin with the same segments but the other accent pattern (Experiment 2).

In addition, we also investigate how bilingual Swedish speakers (simultaneous bilinguals) who grew up with Swedish and one other nontonal language spoken in the family deal with tonal contrasts. This is interesting because for these speakers tonal accent is a meaningful cue in one of their languages but not in the other.

Previous work has addressed the development of bilinguals' discrimination of segmental contrasts that are specific to one of their languages. Bosch & Sebastián-Gallés (2003) reported for Spanish-Catalan bilingual infants that their performance in discriminating a Catalan contrast first declines (by 8 months) and then improves again (by 12 months). The authors attributed this time course, i.e. the renewed discrimination ability after initial decline, to the necessity to represent the contrast during word learning. This is particularly interesting in the present context because, as discussed, tonal word accent in Swedish has very low functional load, implying that such pressure from the lexicon might not exist.

One possibility is that over time bilinguals simply learn to separate their languages perfectly and are thus able to deal with each phonological system independently, processing each language just like monolinguals – including the use of tone in just one of their languages. Another possibility is, however, that the challenge is not so easily dealt with, and the phonological system of bilinguals shows differences compared to monolinguals. If the acquisition of accent regularities in Swedish bilinguals are compromised due to the additional input language, we would expect the sensitivity to accent mispronunciations to be reduced, and therefore predict differences in priming.

Before we introduce our paradigm in detail we review below the existing literature on both the role of suprasegmental features in lexical access and tone processing in bilinguals. We then provide an overview of the nature of tonal word accent in Swedish before reporting our experiments in detail.

Suprasegmental information in lexical access

Suprasegmental information can also play a role for lexical access in non-tonal languages, e.g. in the case of word stress. For instance, stress sensitive, restrictive priming is found in Spanish (Soto-Faraco, Sebastián-Gallés, & Cutler, 2001). Even in languages where stress is correlated with vowel quality, e.g. English (Cooper, Cutler, & Wales, 2002) and Dutch (Cutler & van Donselaar, 2001), native speakers have been shown to exploit stress information in carefully controlled experimental procedures (Cooper et al., 2002).

As mentioned above, in "fully tonal" languages such as Mandarin and Cantonese tone affects the acceptance of a lexeme as a word (Fox & Unkefer, 1985) and restricts priming (Lee, 2007). Cutler & Chen (1997) reported that mispronunciations were more likely to be accepted as words if only tonal contours differed between the mispronunciation and an existing word, indicating that processing for tonal differences is slower or less relevant than segmental differences. However, Malins & Joannisse (2010) used a visual world eye tracking paradigm to investigate the time course of activation for a target item vs. competitors differing either in tone or in segmental information. They found that eye movements to the two types of competitors occurred at a similar time scale, indicating that segmental and tonal information play equal roles in lexical access.

By comparison to the fully tonal Asian languages, Japanese is more relevant for our present work on Swedish because the tonal accent system is a pitch accent system, bearing a similar amount of functional load (cf. Hyman, 2009). Using priming with words that were segmentally identical to the target, and had accurate vs. mispronounced pitch accent, Cutler and Otake (1999) showed that for Japanese listeners pitch accent is an important component in spoken word recognition, and similar results were obtained by Sekiguchi and Nakajima (1999).

Processing of suprasegmental information in bilinguals

Previous work indicates that, as far as the processing of suprasegmental features is concerned, bilinguals can possess monolingual-like skills. There is, however, evidence that this may be restricted to simultaneous bilinguals for whom the target language is dominant, as shown for stress patterns in Spanish by Dupoux, Peperkamp, & Sebastián-Gallés (2010). Their simultaneous Spanish/French bilingual subjects showed monolingual-like performance if Spanish was their dominant language, but French-dominant subjects' responses resembled those of late bilinguals (native French speakers).

While Dupoux et al.'s study involved a speeded lexical decision task, a large portion of the previous work on prosodic processing in bilinguals focussed on auditory perception or discrimination skills (e.g., Abboub, Bijeljac-Babic, Serres, & Nazzi, 2015; Bijeljac-Babic, Höhle, & Nazzi, 2016; Zembrzuski et al., 2020). Work on bilinguals' tone processing has typically targeted speakers of Asian tonal languages (e.g., Yang and Liu, 2012; Huang &

Johnson, 2011; So & Best; 2010), where high functional load makes accurate tone processing a necessity for comprehension. Native speakers of tonal languages show discrimination differences compared to native speakers of non-tonal languages (e.g. Yang & Liu, 2012; Huang & Johnson, 2011) and whether or not tested tone categories correspond to native tone categories also affects tone identification (So & Best, 2010). Investigating tone in a Scandinavian language (Norwegian), van Dommelen & Husby (2009) found that Mandarin native speakers, i.e. subjects with experience in a tonal language, had better performance on an AXB identification test, compared to German natives, who had no experience with tone. Gosselke Berthelsen et al. (2020) presented a pseudo-word learning study which shows consistent results – learners with a tonal native language (Swedish) process tone differently in novel words compared to learners from a non-tonal background (German). There is also previous work investigating the question of whether suprasegmental features used in the native language affect processing in a second, nontonal language. Qin et al. (2017), for example reported native Mandarin listeners, who use pitch to discriminate tone in their native language, relying more on pitch than duration to discriminate stress in English, compared to English native speakers.

Some previous work has addressed lexical access, which means it is more relevant for the present study. Looking at acquisition, Singh & Foong (2012) found that by 11 months, Mandarin/English bilingual infants' recognition of words was restricted to the correct tone in Mandarin, but independent of pitch for English. Working with adults, Wiener & Ito (2015) conducted a visual world experiment with monodialectal Mandarin native (Beijing) and bidialectal speakers, investigating the time course of target identification among a set of phonologically related distracters. All participant groups showed facilitation of target selection in manual responses for low frequency syllables with high probability tones. Eye movements, however, revealed that bidialectal speakers were in fact less sensitive to tonal information. Other studies have addressed language co-activation in bilingual speakers (e.g., Shook and Marian, 2015; Ortega-Llebaria, Nemogá, & Presson et al., 2017), and found that tonal contours presented on an English stimulus can affect processing in Chinese, e.g. providing an advantage in case of a tonal match.

Taken together, these results provide converging evidence that experience with a tonal language affects intonation processing in another language, and may cause interference or decreased sensitivity if this other language has its own tonal inventory.

The nature of Swedish tonal accent

In Swedish there are two contrastive pitch accent patterns, known as Accent 1 and Accent 2. In Central Swedish (Stockholm), acoustically Accent 1 consists of a high tone on the main stressed syllable followed by a low tone, i.e. exhibits one peak, while Accent 2 exhibits two peaks, one on the first and one on the second syllable in disyllabic words. Examples of pitch contours for both patterns are depicted in Figure 1. Monosyllables always carry Accent 1, e.g. *katt*$_1$ ‘cat’, *bil*$_1$ ‘car’, *häst*$_1$ ‘horse’ (in a monosyllable, the Accent 2 pattern, which involves a peak on the second syllable, cannot be realised). The most noticeable contrast is within disyllabic words with initial stress, on which we focus in our paper, e.g. *paddel*$_1$ ‘paddle’ vs. *padda*$_2$ ‘toad’ (see Figure 1). Even within disyllabic words, the distribution of the two accents differs. The vast majority of mono-morphemic disyllabic words are Accent 2 and they constitute a more homogenous set than Accent 1 words, behaving like a "default" in terms of distribution (cf. Lahiri, Wetterlin, & Jönsson-Steiner, 2005; also see Gårding, 1977 (p. 10), who notes that traditionally Accent 1 is associated with monosyllables, whereas Accent 2 is associated with polysyllables). Disyllabic Accent 1 words include words of specific semantic classes, such as days of the week, months, nautical terms, and names of berries, words ending in *-is,* and loans. As discussed above, the functional load of tonal accent is considered to be low in Swedish, which is emphasized by the fact that there are local varieties that do not use tonal accent contrastively. Yet it is considered a distinctive feature of Swedish prosody.

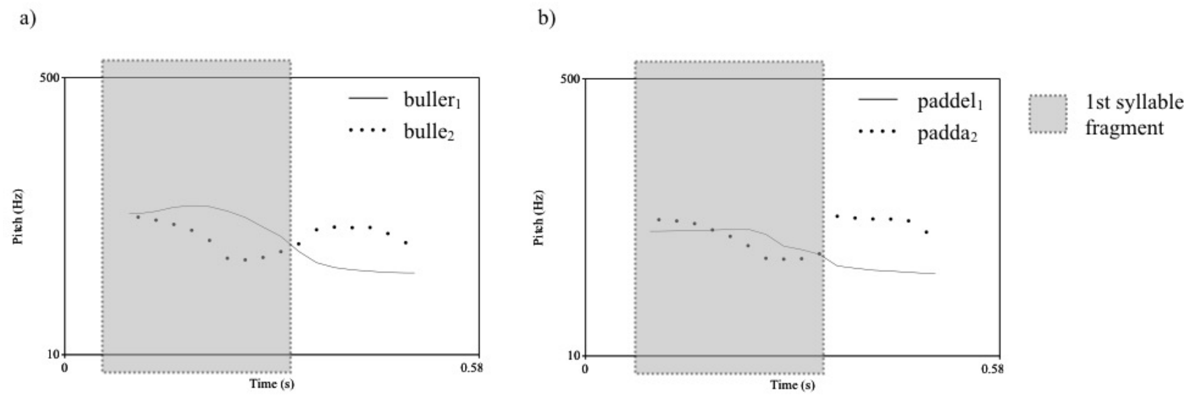

Figure 1. Tonal contours (extracted with Praat; Boersma & Weenink, 2016) for two pairs of accent competitors with segmentally identical first syllable (shaded part) but opposite accent patterns, a) with a sonorant medial consonant and b) with an obstruent medial consonant.

Despite the low functional load of tonal accent in Swedish, Felder et al. (2009) provided experimental evidence that accent plays a role in word identification. Using a fragment completion paradigm they demonstrated that native speakers can reliably discriminate monomorphemic minimal pairs on the basis of accent information contained in just the first syllable of a word. The experimental task involved deciding between two potential target completions of the presented fragment, thus it might have been possible to solve the task without semantic-level activation, i.e. solely using phonological codes.

Several studies by Roll and colleagues have investigated the predictive roll of word accent in complex words (Roll, Horne, & Lindgren, 2010; Söderström, Horne, & Roll, 2017). These studies have found that listeners can predict the correct suffix by just hearing the first stressed syllable of a complex word. For instance, Roll et al. (2010) used an ERP paradigm to demonstrate that the accent of the first syllable is used to predict the second. They investigated monosyllables, where there is a plural alternation (e.g. *mink*$_1$ *sg.* 'mink'*, mink-en*$_1$ *sg.* 'the mink' – *mink-ar*$_2$ *pl.* 'minks') and showed that after hearing the first syllable *mink*$_1$- subjects showed evidence for reanalysis in an increased P600 when presented with the unexpected suffix {–ar}, which is a morphologically valid suffix, but presents an accent mismatch. Importantly subjects did not show an increase in the N400, which indicates morphological mismatch, and which is elicited for the morphologically invalid suffix {-or}. Söderström, Horne, Frid and Roll (2016) attributed differences in the degree to which the first syllable allows predicting upcoming material to the number of competitors in the lexicon. Schremm, Novén, Horne, Söderström, van Westen, & Roll (2018) presented structural brain imaging data relating the degree of individual participants' use of the accent pattern on the stem to predict the suffix to cortical thickness in the left planum temporale. According to Roll and colleagues word accent is therefore induced by the suffix, rather than being stored in the lexicon. There are, however, many simplex/monomorphemic words in Swedish that exhibit accent opposition without having a suffix that could induce it (e.g. Accent 1 words like *taxi*$_1$ 'taxi', *curry*$_1$ 'curry', and disyllabic compounds that are Accent 2, such as *fotboll*$_2$ 'football'). What is clear from the work by Roll and colleagues is that given the pitch contour on the first syllable, native speakers have expectations for the second syllable.

Schremm, Söderström, Horne & Roll (2016) further examined the processing of Swedish tonal word accent in L2 speakers. They presented 3rd Person singular verb forms that show an accent alternation between present tense and past tense (e.g. *han läk-er*$_1$ 'he heals' vs. *han läk-te*$_2$ 'he healed'). In their study, the first syllable of such a verb form was followed either by a suffix that was consistent with the accent pattern presented in the first syllable (i.e. *läk*$_1$

followed by *-er*, or *läk*$_2$ followed by *-te*) or with an inconsistent syllable (i.e. *läk*$_1$ followed by *-te*, or *läk*$_2$ followed by *-er*). Similar to native speakers, L2 speakers showed an increase in response times for the inconsistent combinations, indicating that L2 speakers of Swedish represent accent patterns, despite their low functional load.

Here we focus specifically on the role of word accent in lexical access. While Felder et al. (2009) showed that native speakers readily use accent in word discrimination, their task may have only tapped into a phonological level of processing. Here, we therefore introduce a semantic priming task which is highly sensitive due to the use of first syllable fragments as primes.

Rationale for the present studies

Our aim was to investigate the role of accent information in lexical access. Swedish allows pairs of words where the first syllables have complete segmental overlap but which differ in accent: *tanke*$_2$ 'thought' ~ *tango*$_1$ 'tango'. We also find instances where the first syllable does not lead to an accent competitor; e.g., the syllable [pɔn$_1$] is part of *ponny*$_1$ 'pony', but there is no noun which begins with *pon-* which has Accent 2.

Experiments 1 and 2 make use of these possibilities in a cross-modal semantic priming task with fragments as primes. Here, the first syllable of an auditory prime (e.g. [pɔn$_1$] for *ponny*$_1$ 'pony') is used to activate a lexical item, as measured by the priming of a semantically related, visually presented target (e.g. *häst* 'horse') in a lexical decision task. If accent plays a role in lexical access, then accent mispronunciations of the prime ([*pɔn$_2$]) should not lead to priming. Experiment 1 uses pairs of words that share the same first syllable in terms of segments, but which differ in word accent. In other words, there is a direct *accent competitor* for each of the prime items. For instance, the "mispronunciation" of [taŋ$_2$] (for *tanke*$_2$ 'thought') is [taŋ$_1$], which is the *correct* pronunciation of the first syllable of the existing *tango*$_1$ 'tango' (see Figure 2). Felder et al. (2009) used such word pairs in a fragment completion task (i.e. without semantic priming), so we already know that native speakers' performance is high with regard to matching the accent pattern they hear to the corresponding lexical entry.

Experiment 2, by contrast, focusses on words where no direct *accent competitor* (in terms of the first syllable) is found in the lexicon, i.e. mispronunciation of the accent in the first syllable, such as [*pɔn$_2$] for *ponny*$_1$, results in a non-existing first syllable (see Figure 3).

We argue that the method of semantic priming with fragments is highly sensitive due to the fragile link between the fragment prime and the semantically related target; a fragment needs to suffice for lexical activation, which in turn needs to activate semantically related words in order for priming of the target to be observed. At the same time this has the chance of providing quite strong evidence for the role of accent, since semantic representations (rather than just phonological) are necessarily involved.

Our second aim was to assess whether lexical accent is used to the same extent as by monolinguals by simultaneous bilingual speakers, who grew up with Swedish and one other (non-tonal) native language. In other words, we are asking whether the acquisition of a low-functional load feature is compromised once it is no longer a universal cue to lexical identity across all of the language input received by the learner.

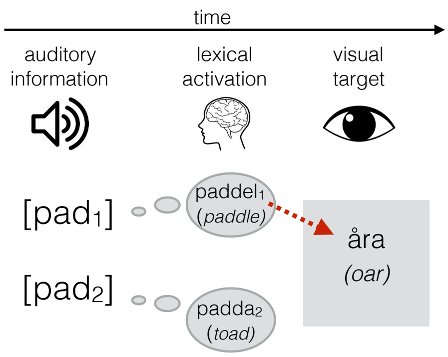

Figure 2. Schematic illustration of design of Experiment 1 with primes and accent competitors: Mispronunciation results in an existing first syllable, leading to lexical competition.

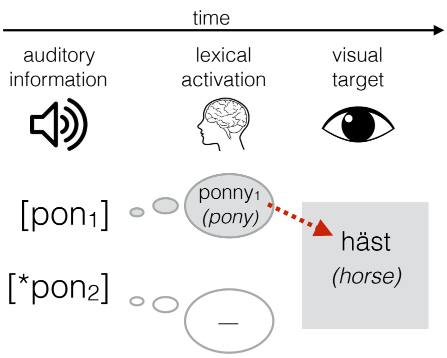

Figure 3. Schematic illustration of the design of Experiment 2 with primes and accent mispronunciations. Here there are no accent competitors, i.e. the fragment with opposing accent pattern is a true mispronunciation.

Experiment 1

Here, prime words were chosen by constructing pairs of accent competitors, i.e. words whose first syllables consist of the same segments, but whose accent differs. Accent mispronunciations of the first syllable fragment therefore result in an existing first syllable and ought to activate a different lexical entry - causing lexical competition. These competing lexical entries should not cause priming of the visually presented target. We hypothesized that having a competitor in the lexicon would force speakers to discriminate precisely between the Accent 1 and Accent 2 word and therefore priming should be restricted to accurate pronunciations. Thus, [taŋ$_1$] should be discriminated from [taŋ$_2$] because they are beginnings of two different words, *tango*$_1$ 'tango' and *tanke*$_2$ 'thought'.

Methods

All experiments reported here were conducted at Stockholm University, Department of Swedish Language and Multilingualism. Furthermore, the stimuli were recorded by a female speaker from that area.

Participants

Fifty-one participants took part in this study (average age 27.3 years). Thirty-six of them were monolingual native speakers of Swedish; i.e., Swedish was the language spoken in the household they grew up in. All of them learned further languages in school and elsewhere. Fifteen participants grew up with Swedish along with one other native language, to which they were exposed to from birth from one parent. These languages were English, Farsi, French, Finnish, German, Icelandic, Serbian, Spanish, Syrian, and Turkish. Participants with other tonal native languages such as Mandarin (N = 3) were excluded, as well as Finnish native speakers who indicated that they had grown up in Finland with a Swedish-speaking parent (N = 3) due to the possibility that they were mainly exposed to Finland Swedish, which does not use tonal accent. All bilingual participants lived and studied or worked in Stockholm at the time of participation. Monolingual participants lived and studied or worked in Stockholm at the time of participation, and all had learned at least one language from school age, but none had heard a second language spoken at home or from earlier on in their life. Participants were recruited via an online participant recruitment system as well as via flyers, posters and facebook groups at Stockholm University.

Stimuli and design

As stimuli we selected sixty-eight pairs of disyllabic nouns beginning with segmentally identical first syllables, but differing in accent. The majority of these were taken from Felder et al. (2009), but we added the following extra pairs: *bingo*$_1$/*binge*$_2$, *facit*$_1$/*fasa*$_2$, *friskis*$_1$/*frissa*$_2$, *harem*$_1$/*hare*$_2$, *iglo*$_1$/*igel*$_2$, *kändis*$_1$/*känsla*$_2$, *avel*$_1$/*avund*$_2$, *vimpel*$_1$/*vimmel*$_2$, *dimmer*$_1$/*dimma*$_2$, *halster*$_1$/ *halma*$_2$, *hagel*$_1$/*hage*$_2$, *joller*$_1$/*jolle*$_2$, *vimmel*$_1$/*vimpel*$_2$ (see Table 4, Online Supplementary Material).

Half of the fragments stemmed from disyllabic Accent 1 primes and the other half from disyllabic Accent 2 primes. Accent 1 and Accent 2 words were approximately matched in terms of frequency. Targets were semantically related to the prime words. All the prime words were recorded by our speaker. After normalising the maximum volume of all recorded words, fragments of the primes were extracted in Praat (Boersma & Weenink, 2018) by cutting out the first syllable of the word. Care was taken to ensure that correct and mispronounced versions were equally long in milliseconds (average length = 331 ms) and the same amount of segmental information was provided in both versions. If one fragment had a shorter vowel compared to the other, then the duration of the shorter one determined the maximum length of both fragments. No artificial lengthening or shortening was used. Fragments were always cut at zero-crossings and at most 2 glottal pulses were removed from vowels when it was necessary to match the duration of pairs. For words with an initial open syllable such as *hagel*$_1$ [ˈhɑːgəl] the syllable ended in the vowel (e.g., [ˈhɑː]), while for words with a medial cluster and an initial closed syllable (e.g., *kändis*$_1$ [ˈɕɛndɪs]), the first syllable ended in the consonant (e.g., [ˈɕɛn]), see Figure 4 for example spectrograms. For some of the paired Accent 1 / Accent 2 words with open first syllable fragments the upcoming medial segments were the same (e.g. *[ha]gel*$_1$ – *[ha]ge*$_2$ with upcoming consonant /g/) and for some the upcoming segments were different (e.g. *[fa]**b**el*$_1$ – *[fa]**d**er*$_2$*, with upcoming consonants /b/ vs. /d/)*. In order to make sure the second case (different upcoming consonants) did not result in words being identified on the basis of co-articulation information present on the fragment due to the following consonant, we conducted a formant analysis on the fragments for all pairs of words used in Experiment 1, which can be found in Appendix A. This showed that differences between fragment pairs with regard to formant frequencies at the fragment offset were not different for pairs with the same vs. different upcoming consonants. We can therefore rule out effects of co-articulation information.

In addition, 136 nonword targets were created and paired with fragments of words not otherwise used in the experiment. Four stimulus sequences were created such that each contained 272 trials. Half the trials contained control words with nonword targets. Of the remaining 136 trials, half consisted of prime-target pairs that were neither semantically nor phonologically related. The remaining pairs contained semantically related prime-target pairs (e.g., *hallon* 'raspberry' – *bär* 'berry') where the prime fragments either had the correct or incorrect tonal accent (34 correct, 34 mispronounced; see Table 1 for an illustration of the design). Each target occurred exactly once during each of the four sequences, meaning that no participant saw a target more than once. The four sequences were assembled such that each related prime-target pair occurred together in two sequences (once with a correct, and once with a mispronounced prime), and separately as part of unrelated prime-target pairs in the remaining two sequences.

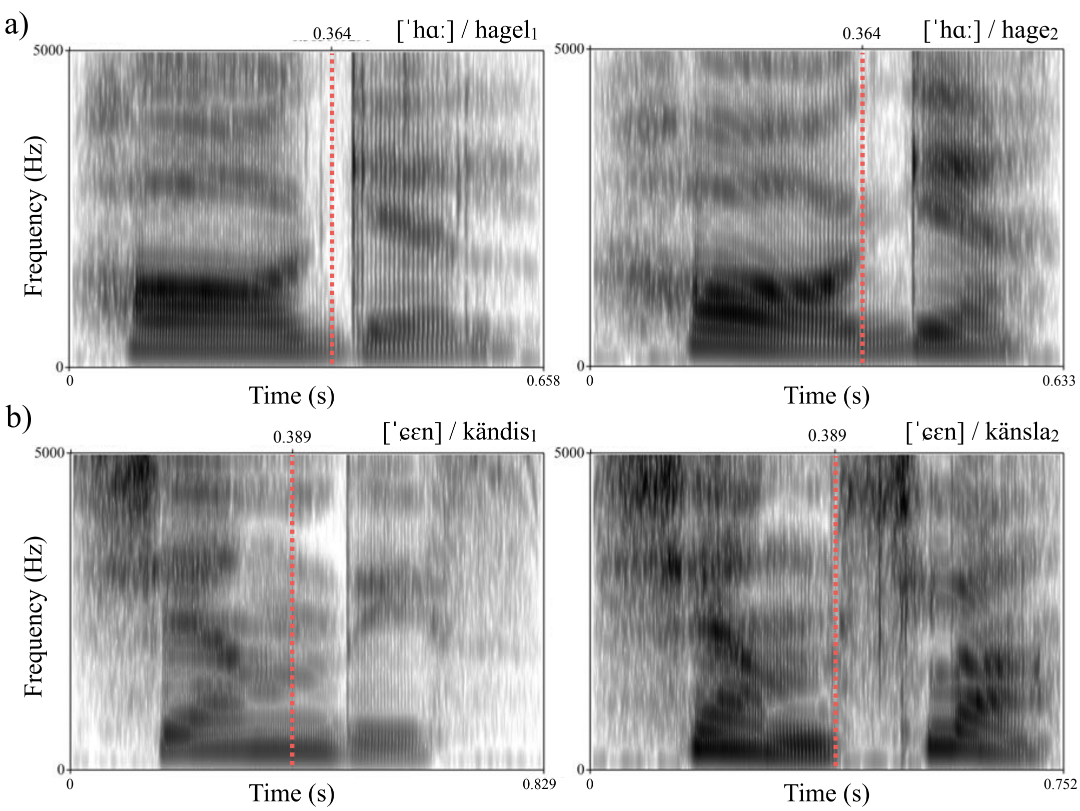

Figure 4. *Spectrograms for Accent 1 and Accent 2 pairs of words with (a) open (hagel$_1$ / hage$_2$) and (b) closed (kändis$_1$/känsla$_2$) first syllable fragments. Dashed red lines indicate the end (cut point) of the first syllable fragment.*

Table 1.

Critical trial types in Experiments 1 and 2.

Pronunciation	Prime type	Correct Accent	Prime	Fragment (auditory)	Target (visual)
Correct	Related	1	*ponny*$_1$ ‘pony’	[pɔn]$_1$	*häst* ‘horse’
Correct	Related	2	*färja*$_2$ ‘ferry’	[fɛr]$_2$	*båt* ‘boat’
Mispronounced	Related	1	**ponny*$_2$ ‘pony’	*[pɔn]$_2$	*häst* ‘horse’
Mispronounced	Related	2	**färja*$_1$ ‘ferry’	*[fɛr]$_1$	*båt* ‘boat’
Correct	Unrelated	1	*pon*$_1$ ‘pony’	[pɔn]$_1$	*snö* ‘snow’
Correct	Unrelated	2	*färja*$_2$ ‘ferry’	[fɛr]$_2$	*fiol* ‘violin’
Mispronounced	Unrelated	1	**ponny*$_2$ ‘pony’	*[pɔn]$_2$	*snö* ‘snow’
Mispronounced	Unrelated	2	**färja*$_1$ ‘ferry’	*[fɛr]$_1$	*fiol* ‘violin’

Procedure

Consent was obtained in writing at the beginning of the session. Up to six participants could be tested simultaneously, each being equipped with individual response boxes and head sets. Words were displayed on a single screen mounted on the wall of the room, in white on a black background. The screen was approximately 2-4 m from individual participants. The room was darkened to ensure maximum visibility of the stimuli, and minimum distraction through the presence of other participants. Recorded prime fragments were presented to individual head sets. The prime-target onset latency was 500 ms. Subjects were instructed to respond as quickly and as accurately as possible only to the visually presented stimuli via the response box with two buttons labelled “yes” (for words) and “no” (for nonwords). Before the start of the real experiment, participants were presented with five practice trials containing words not used in the experiment to enable them to become familiar with the procedure.

Stimulus selection

Latencies for correct response times were calculated for each target word. Outliers (response time larger than two standard deviations above the mean) were removed from the data. Since no priming norms were available for the initial stimulus selection, we first analysed only the correctly-pronounced trials from monolinguals to establish which prime-target pairs were effective in priming. Mispronounced trials and data from bilinguals were not included in this selection step. See Table 4 (Online Supplementary Material) for a list of all stimuli and difference in mean RT (Unrelated – Related).

We calculated the mean difference in response time for each word when it occurred in an unrelated trial and when it occurred in a related trial, for monolingual participants. The prime was judged effective if this difference was larger than 0 ms. We identified 57 effective prime-target pairs. Reaction times for this set of items are shown in Figure 5. The average response times were M = 547 ms (SE = 2.9 ms) for related trials, and M = 563.3 ms (SE = 2.95 ms) for unrelated trials. The difference between related and unrelated trials is significant (a linear mixed effects model with fixed effect Relatedness, and crossed random effects of participants and items on intercepts, as well as random effects of participants on slopes, yields the following for the unrelated items with respect to the "related" baseline: estimate = 16.37, SE = 3.5, t(43.7) = 4.66, p < .0001).

Results

We then proceeded to examine the mispronunciation effect on the related trials only (a mispronunciation effect on unrelated trials is not expected). Related accurately pronounced trials and related mispronounced trials from both monolinguals and bilinguals were included in these models.

We fitted linear mixed-effects models to the data from related trials using the lme4 package in R (Bates, Maechler, Bolker, & Walker, 2015). We began with a base model that contained only random effects of *Participants* and *Targets* on intercepts, and then added the fixed factor *Pronunciation* (Correct, Mispronounced). We also added random effects of *Participants* on the slopes. A model comparison showed that this was a significant improvement in model fit (χ^2 (3)= 21.98, p < .0001). We then fitted a model adding the fixed factor *Accurate Prime Accent* (as well as random effects of *Participants* on intercepts and slopes). This did not improve model fit further (χ^2 (4) = 1.79, p > .78), nor did the addition of an interaction term (χ^2 (5) = 10.26, p = .07). Adding the factor *Bilingual Status* (χ^2 (1) = 0.13, p > .71), as well as the corresponding interaction term (χ^2 (3) = 2.23, p > .52) did not improve the model fit either. The best-fitting model was therefore the model with a fixed effect of *Pronunciation* and random effects of *Participants* on intercepts and slopes, and random effects of *Targets* on intercepts. Here, the main effect of *Pronunciation* was significant ($t(462.868)= 4.451$, $p < .0001$, see Table 2), with accurate pronunciations estimated to lead to 21 ms faster responses compared to mispronunciations.

Table 2.

Summary of fixed effects in the best-fitting model for Experiment 1 obtained with lme4 and lmerTest (Kuznetsova, Brockhoff, and Christensen, 2017) packages. Effect of Pronunciation estimated with respect to base "Correct".

	Estimate	Std. Error	df	*t*	*p*
Intercept	539.79	10.46	62.073	51.609	<.0001
Pronunciation (Mispronounced-Correct)	21.29	4.783	462.868	4.451	<.0001

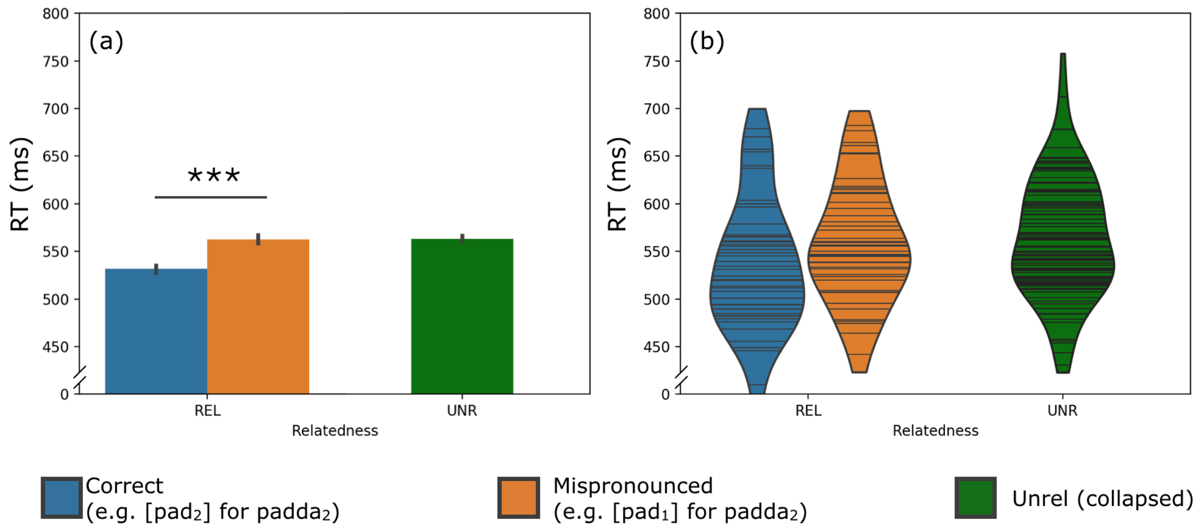

Figure 5. Experiment 1 results (primes with accent competitors) for accurately pronounced related trials, mispronounced related trials and unrelated trials. (a) Bar charts indicating mean response times (unaggregated). Error bars show standard errors of the mean. Asterisks indicate contrasts significant in the model (p<.0001). (b) Violin plots showing the distribution of participant means (horizontal lines) for each condition.

Discussion

Experiment 1 results indicate that priming is restricted to accurate accent pronunciation with the present set of words. This seems to be the case to an equal extent for Accent 1 and Accent 2 words. This is a particularly strong result given that primes consisted of only the first syllable fragment. Differences between accurate and mispronounced items are clearly noticeable even in the first stages of word processing, before the whole word has been heard. If the correct fragment is heard, a lexical entry is activated and this primes semantically related items. By contrast, if the fragment has the wrong accent, no such activation occurs and hence no priming of semantically related items follows.

We found no evidence of differences between monolingual and bilingual speakers. Both of our participant groups identify themselves as native speakers, but their language input during acquisition was different in the sense that the bilinguals received an additional, non-tonal language. Our hypothesis was that this might complicate the process of extracting tonal accent as a feature relevant for lexical access and we might therefore see compromised performance. This was, however, not the case. Monolinguals and bilinguals had similar effects of relatedness and pronunciation. There is therefore no evidence that exposure to a second, non-tonal, native language disrupts the extraction of word accent as a feature relevant in lexical access.

However, Experiment 1 represents a special case: here, we used items that were specifically selected because they have an accent competitor, i.e. there is a noun in the lexicon that has a first syllable consisting of the same segments but with the opposite accent pattern. It is therefore possible that our results reflect not a general sensitivity to accent mispronunciation but are specifically due to competition effects. Perhaps priming is only absent for accent mispronunciation where such a mispronunciation results in an existing first syllable, as was the case in Experiment 1. It is possible that in general listeners would handle accent mispronunciations and still show semantic priming, but here, because there is pressure from the lexicon to distinguish between two segmentally identical syllables, e.g. [pad_1] and [pad_2] for *paddel*$_1$ and *padda*$_2$, only an accurately pronounced first syllable segment narrows down lexical activation enough to achieve priming. We therefore conducted a second experiment, this time selecting Accent 1 and Accent 2 items on the basis that they had no accent competitor, i.e. mispronunciation of the first syllable with the opposing tonal accent results in a nonexistent form.

Experiment 2

Methods

Participants

Fifty-four participants took part in the study (average age: 26.9 years). Thirty-nine of them were monolingual native speakers of Swedish (i.e. only Swedish was spoken in the household they grew up in; all of them learned further languages in school and elsewhere), and 15 grew up with Swedish and at least one other native language spoken by one of their parents (spoken to them from birth). These participants listed English, French, Finnish, German, Greek, Serbian, Spanish, Syrian, Swahili, and Turkish as native languages (besides Swedish). As in Experiment 1, participants with other tonal native languages such as Mandarin (N = 3) were excluded, as well as Finnish native speakers who had grown up in Finland with a Swedish-speaking parent (N = 3). The monolinguals had learned at least one other language in school, but none of them had heard a language other than Swedish spoken at home or from an earlier age. All participants lived and studied or worked in Stockholm at the time of participation. Some participants took part in both of our experiments, but these were conducted on different days. Some participants first took part in Experiment 1 and some first in Experiment 2.

Stimuli, design and procedure

Since most participants also took part in Experiment 1, we ensured that there was no overlap in primes and targets used across the two experiments.

A total of 128 prime-target pairs served as the stimuli in this experiment (see Table 5, Online Supplementary Material). All prime words were chosen such that they did not have an accent competitor, i.e. an existing lexical entry (noun) with a segmentally identical first syllable but exhibiting the other accent pattern was not found. For example the fragment [ˈpɔn$_1$] from *ponny$_1$* does not have an accent competitor [ˈpɔn$_2$] in the lexicon. In other words, accent mispronunciations (of the first syllable fragment) should not produce an accurate match in the lexicon. Half of the fragments stemmed from disyllabic Accent 1 primes and the other half from disyllabic Accent 2 primes~~.~~ Accent 1 and Accent 2 words were approximately matched in terms of frequency. Targets were semantically related to the prime words. All the prime words were recorded by our speaker who produced one accurately pronounced and one mispronounced version. The mispronunciation was kept as similar as possible to the correct version, but pronounced with the opposite tonal contour. Stimuli were preprocessed as described in Experiment 1.

Four stimulus sequences were created such that each contained 256 trials. Half the trials contained control words with nonword targets. Of the remaining 128 trials, half consisted of prime-target pairs that were neither semantically nor phonologically related. The remaining pairs contained semantically related prime-target pairs (e.g., *färja* ‘ferry’ – *båt* ‘boat’) where the prime fragments either had the correct or incorrect tonal accent (32 correct, 32 mispronounced; see Table 1 for an illustration of the design). Each target occurred exactly once during each of the four sequences, meaning that no participant saw a target more than once. The four sequences were assembled such that each related prime-target pair occurred together in two sequences (once with a correct, and once with a mispronounced prime), and separately as part of unrelated prime-target pairs in the remaining two sequences.

The procedure was as described for Experiment 1.

Stimulus selection

Latencies for correct response times were calculated for each target word. As in Experiment 1, in the absence of available priming norms we first established which primes were successful by analysing only the correctly pronounced trials from monolinguals. As above, data from bilinguals, and data from mispronounced trials were not included in this step. We calculated the mean difference in response time for each word when it occurred in an unrelated trial and when it occurred in a (correctly pronounced) related trial, for monolingual participants, with a prime being judged as effective if this difference was larger than 0 ms. See Table 5 (Online Supplementary Material) for a list of all stimuli and corresponding difference in mean RT (Unrelated – Related). We identified a set of 85 targets for which monolinguals benefited from a semantically related, accurately pronounced prime fragment (43 of these were Accent 1 items and 42 were Accent 2 items). Only these stimuli were selected for further analyses. This subset selection implies that priming in accurately pronounced trials is trivially successful (Related trials, including correct and mispronounced primes: $M = 555$ ms, $SE = 2.9$; Unrelated trials: $M = 583$ ms, $SE = 2.9$), the subsequent analysis focused on examining whether there was a mispronunciation effect across semantically related trials.

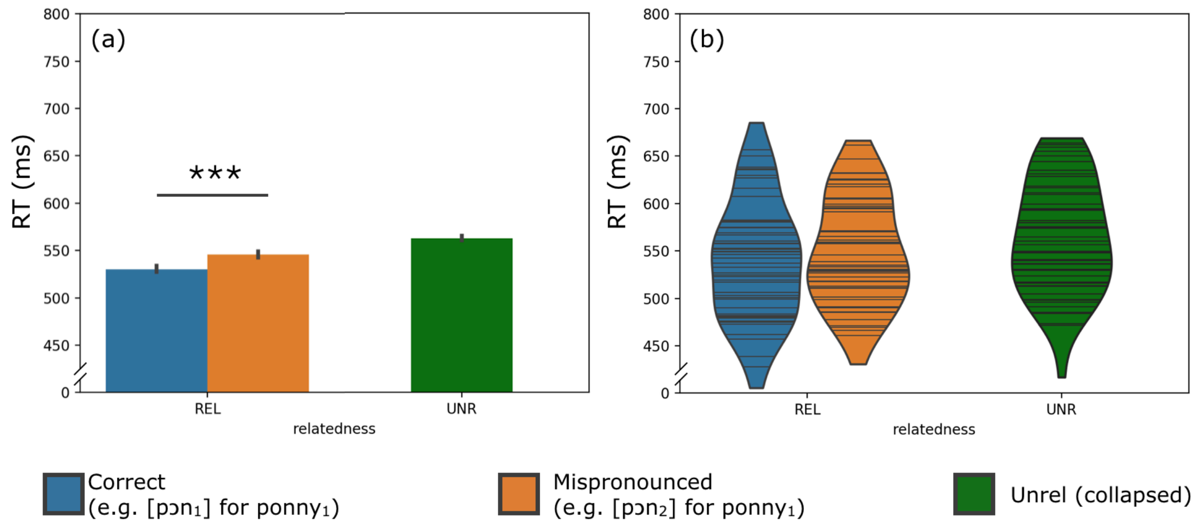

Figure 6. Experiment 2 results (primes without accent competitors) for accurately pronounced related trials, mispronounced related trials, and unrelated trials. (a) Barplot showing mean response times (unaggregated). Error bars indicate standard errors of the mean. Asterisks indicate contrasts significant in the model (p=.0003). (b) Violin plot showing the distribution of participant means (horizontal lines) for each condition.

Results

Average response times are shown in Figure 6. As in Experiment 1, we modelled the response time data from related trials with a linear mixed effects model (as before, a mispronunciation effect is not expected for unrelated trials, so these were not included in the model). We first constructed a base model with only random effects of *Participants* and *Targets* on intercepts. We then added fixed factors in a stepwise fashion, starting with *Pronunciation* (correct, mispronounced) and then *Accurate Prime Accent* (1, 2), and *Bilingual Status* (monolingual, bilingual). For *Pronunciation* and *Accurate Prime Accent* we also added random effects of *Participants* to the slopes. Model comparisons showed that adding *Pronunciation* to the base model improved the model fit (χ^2 (3) = 13.126, p = 0.004). Adding the effect of *Accurate Prime Accent* did not further improve the model fit (χ^2(4)=6.15, p = 0.188), nor did adding an interaction with *Pronunciation* (χ^2 (5) = 3.52, p > 0.61. The addition of *Bilingual Status* did not improve the model fit further (χ^2(1)=0.001, p > 0.97), nor did the interaction of *Bilingual Status* with *Pronunciation* and *Accent* (χ^2(3)=1.84, p > 0.6). The best model fit was therefore achieved with a model that included only the fixed factor of *Pronunciation*. The model, summarised in Table 3, exhibits a main effect of *Pronunciation*, with accurately pronounced items an estimated 17 ms faster than mispronunciations.

Table 3.

Summary of fixed effects in the best-fitting model for Experiment 2 (obtained with package lmerTest), effect estimated with respect to base Pronunciation = correct

	Estimate	Std. Error	df	*t*	*p*
Intercept	553.062	11.036	76.788	50.12	<.0001 ***
Pronunciation (Mispronounced-Correct)	16.968	4.701	1170.346	3.61	0.0003 ***

Discussion

The results from Experiment 2 show that lexical word accent restricts lexical access even in the absence of direct lexical competition. Clearly, for both monolingual and bilingual speakers primes are only efficient when pronounced accurately. This confirms our findings from Experiment 1 and further supports our conclusion that full activation of the prime lexeme, to the extent that it allows propagation of activation to semantically related items, only occurs when the correct accent is heard. Moreover, Experiment 2 demonstrates that lexical word accent appears to be necessary for this to occur even in cases where the first syllable segments alone are in fact sufficient to identify the lexical entry, i.e. the accent patterns as such are in fact redundant. Consistent with the results from Experiment 1 there is no evidence that bilinguals deal with accent patterns differently compared to monolinguals.

Overall Discussion

Swedish utilises a tonal accent contrast in its phonological system, with contrasting patterns labelled Accent 1 and 2. This contrast, however, has a lower functional load than regular phonological segmental contrasts. Not only do very few minimal pairs with accent contrast exist – in addition there are a few Swedish dialects which do not make use of tonal marking. Nevertheless, for Standard Swedish, lexical word accent is contrastive and does distinguish words. The question we asked is the following: since the accent contrast appears not to have a high functional load but is nevertheless in use in the majority of Swedish dialects, is there a difference in the way monolinguals and bilinguals are sensitive to the contrast? Since bilinguals cope with two languages, would they perceive the accent contrast as 'unnecessary' and 'unimportant'? Or is the contrast so well entrenched into the fundamental phonological system of the language that bilinguals and monolinguals will be sensitive to any deviation in accent pronunciation?

To examine the strength of the accent contrast in the phonological system, we have presented two cross-modal fragment semantic priming experiments to investigate how monolinguals and bilinguals deal with tonal word accent during lexical access. Experiment 1 employed pairs of words as primes which form accent competitors, i.e. their first syllables are segmentally identical but require the opposite accent patterns. Experiment 2 used words for which there is not an accent competitor in the lexicon (as far as the first syllable fragment is concerned). Our results indicated that tonal word accent plays a critical role in lexical access. Importantly, our study demonstrates that accent patterns are used in a situation that is only temporarily ambiguous. Hearing the second syllable resolves the ambiguity at the level of fragments. In other words, even if accent were to be ignored completely, lexical identification would be possible for these words in a real life (non-fragment-priming) scenario. Yet, our participants exhibited restricted priming, indicating that these mechanisms are at work despite providing only a temporary advantage. The implication is that tonal word accent, though of low functional load in its theoretical definition, is useful in Swedish: it contributes to the speed of lexical identification.

Finally, we found no difference between the performance of monolingual native speakers and that of simultaneous bilingual speakers, who in addition to Swedish were exposed to one other, non-tonal language while growing up. The fact that performance is highly similar for these two groups indicates that bilinguals are able to extract and encode contrastive tonal accent in spite of a large part of their input not using tone. At least as far as semantic priming is concerned, and the level of representation this task taps into, having a second native language in which tone does not play a role does not seem to interfere with the ability to use tonal word accent during lexical access in Swedish.

Note, however, that the bilinguals in our samples all lived and studied/worked in Stockholm. It is an open question whether performance on these priming tasks would be less robust in bilingual speakers for whom Swedish is not dominant – e.g. simultaneous bilinguals with a Swedish speaking parent but growing up outside Sweden, or late bilinguals (cf. Dupoux et al., 2010).

One aspect worth discussing is that both with and without competition, the priming effect seems to tap into the surface level of representation, rather than the underlying phonological level. Different linguistic analyses of the Accent 1/2 distinction in Swedish have been put forward. Whereas Bruce (Bruce, 1977; Gussenhoven & Bruce, 1999) suggested that both accents are specified in the lexicon, Riad (1998; 2003; 2014) adopted a view of Accent 2 as lexically specified due to its phonetic complexity, assuming instead that Accent 1 is assigned post-lexically. Lahiri et al. (2005) by contrast argued that the most parsimonious analysis treats Accent 2 as a default, i.e. not lexically specified, whereas Accent 1 words must be specified in the lexicon as bearing Accent 1 (also cf. Wetterlin, Lahiri, & Jönsson-Steiner, 2007; Wetterlin, 2010). The latter two views suggest that there should be an asymmetry in processing of Accent 1 and Accent 2. Mispronunciations lack a surface feature, but according to Lahiri et al. (2005), this is not specified in the lexicon for Accent 2 items. We would therefore expect the relevant lexical entry to be activated regardless of the erroneous accent pattern, i.e. mispronunciation should not matter as much as for Accent 1 words. Our current studies do not exhibit strong asymmetries between the two accent types. It is, however, plausible that the differences in lexical specification are just obscured in the complex two-stage process of activating a prime through fragment presentation and activation of semantically related items. After all, the link between prime fragment and (semantically related, i.e. one step removed) target item is particularly fragile in this case.

The crucial finding here is that Swedish tonal word accent represents a vital component of the speech signal that is clearly used in lexical access. Despite the low functional load of tonal word accents in Swedish, even bilingual speakers, for whom the signal-to-noise ratio with regard to this feature during acquisition is even less favourable for extraction, show priming effects restricted to accurate accent pronunciation. This demonstrates that tonal word accent in Swedish has an important status that is not reflected in its functional load. The accent contrast if firmly entrenched in the phonological system of Standard Swedish and all other dialects which make use of it.

Acknowledgements

We wish to thank Tomas Riad and the Institute for Swedish and Multilingualism at the University of Stockholm for kindly hosting us for data collection, Gunnar Norrman for technical support in Stockholm, Colin Brooks for technical support in Oxford, Clara Palm for research assistance, Sara Myrberg for help recording stimuli, Irina Lepădatu, Janette Chow and Jelena Sučević for help with stimulus preprocessing.

Statement of Ethics

The work in these experiments was carried out in accordance with The Code of Ethics of the World Medical Association (Declaration of Helsinki). The work was approved by the University of Oxford's Medical Sciences Division Ethics Committee under the project "MS-IDREC-C1-2015-078: The representation of word sounds in the mental lexicon". All participants provided written informed consent prior to taking part in the study.

Conflict of Interest Statement

The authors have no conflicts of interest to declare.

Funding Sources

This research was funded by ERC Project FP7-IST-269670 to Aditi Lahiri.

Author Contributions

NA, AW and AL designed the research. NA and AW conducted the research. NA analysed the reaction time data. AW carried out detailed phonetic analyses of the stimulus material. NA, AW and AL wrote the paper.

Appendix

Appendix A. Formant analysis for stimuli used in Experiment 1

Pairs of items for which the first syllable fragments serve as accent competitors sometimes contained different medial consonants. In order to make sure that this did not lead to co-articulation information at the end of the fragment, which could have allowed participants to discriminate between the competing items, we conducted a formant analysis on the fragments. Here we provide measurements for formants F1, F2 and F3, taken at the final three glottal pulses of the fragment using Praat. The values from the last three formant measurements were averaged to yield one mean F1 value, one mean F2 value and one mean F3.

For each pair of Accent 1 and matched Accent 2 items (e.g. *fabel*/*fader*), we calculated the difference between the respective F1 means, F2 means and F3 means. We then inspected the distribution of these paired distances across pairs of items where the continuation after the fragment cut consisted of:

(a) the same segment, e.g. *[ha]gel – [ha]ge*, with upcoming consonant /g/

(b) of different segments, e.g. *[fa]**b**el – [fa]**d**er*, with upcoming consonants /b/ vs. /d/

Two-sample t-tests (Welch) were also conducted to compare the distributions, revealing no signficant differences (F1: $t(32.5) = -0.553$, $p = 0.584$; F2: $t(53.3) = 1.242$, $p = 0.220$; F3: $t(33.28) = -1.233$, $p = 0.226$). We therefore concluded that co-articulation information could not have caused participants to respond on information other than tonal accent in Experiment 1.

Figure 7 shows density plots for the distribution of the differences for F1, F2, and F3 respectively.

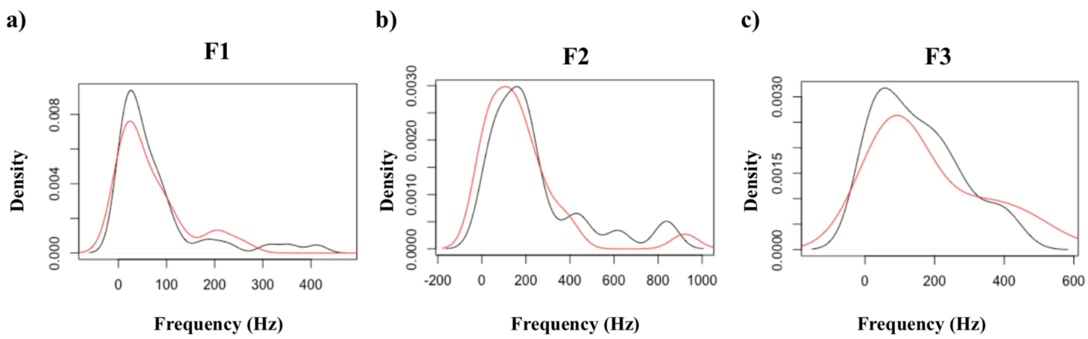

***Figure** 7. Density plots for distribution of differences between measurements for matched Accent 1/Accent 2 pairs in Experiment 1 a) F1, b) F2, c) F3. Distributions for pairs with the same medial segment are plotted in red, distributions for pairs with different medial segments are plotted in black.*

References

Abboub, N., Bijeljac-Babic, R., Serres, J., & Nazzi, T. (2015). On the importance of being bilingual: Word stress processing in a context of segmental variability. *Journal of Experimental Child Psychology*, *132*, 111–120.

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*, *67*(1), 1-48.

Bijeljac-Babic, R., Höhle, B., & Nazzi, T. (2016). Early Prosodic Acquisition in Bilingual Infants: The Case of the Perceptual Trochaic Bias. *Frontiers in Psychology*, *7*(February), 1–8. https://doi.org/10.3389/fpsyg.2016.00210

Boersma, P. & Weenink, D. (2018). Praat: doing phonetics by computer [Computer program]. Version 6.0.37, retrieved 3 February 2018 from http://www.praat.org/

Bosch, L., & Sebastián-Gallés, N. (2003). Simultaneous Bilingualism and the Perception of a Language-Specific Vowel Contrast in the First Year of Life. *Language and Speech*, *46*(2–3), 217–243.

Bruce, G. (1977). *Swedish Word Accents in Sentence Perspective*. Gleerup.

Cooper, N., Cutler, A., & Wales, R. (2002). Constraints of Lexical Stress on Lexical Access in English : Evidence from Native and Non-native Listeners, *45*(3), 207–228.

Cutler, A., & Chen, H. C. (1997). Lexical tone in Cantonese spoken-word processing. *Perception & Psychophysics*, *59*(2), 165–79.

Cutler, A., & Otake, T. (1999). Pitch accent in spoken-word recognition in Japanese. *The Journal of the Acoustical Society of America*, *105*(3), 1877–1888.

Cutler, A., & van Donselaar, W. (2001). Voornaam is not (really) a homophone: Lexical prosody and lexical access in Dutch. *Language and Speech*, *44*(2), 171–195.

Elert, C.-C. (1972). Tonality in Swedish: Rules and a list of minimal pairs. In E. S. Firchow, K. Grimstad, N. Hasselmo, & W. A. O'Neil (Eds.), *Studies for Einar Haugen* (pp. 151-173). Mouton de Gruyter.

Elert, C.-C. (1981). *Ljud och ord i svenskan 2*. Almqvist & Wiksell.

Felder, V., Jönsson-Steiner, E., Eulitz, C., & Lahiri, A. (2009). Asymmetric processing of lexical tonal contrast in Swedish. *Attention, Perception & Psychophysics*, *71*(8), 1890–1899.

Fox, R. A., & Unkefer, J. (1985). The effect of lexical status on the perception of tone. *Journal of Chinese Linguistics,13*(1), 69-90.

Gårding, E. (1977). *The Scandinavian Word Accents.* Lund, Sweden: Gleerup

Gosselke Berthelsen, S., Horne, M., Shtyrov, Y., & Roll, M. (2020). Different neural mechanisms for rapid acquisition of words with grammatical tone in learners from tonal and non-tonal backgrounds: ERP evidence. *Brain Research*, *1729*, 146614.

Gussenhoven, C. (2004). *The Phonology of Tone and Intonation.* Cambridge University Press.

Gussenhoven, C., & Bruce, G. (1999). Word prosody and intonation. In van der Hulst (ed.), *Empirical approaches to language typology*, 233-272. Mouton De Gruyter.

Huang, T., & Johnson, K. (2011). Language specificity in speech perception: Perception of mandarin tones by native and nonnative listeners. *Phonetica*, *67*(4), 243–267.

Hyman, L. M. (2009). How (not) to do phonological typology: the case of pitch-accent. *Language Sciences*, *31*(2–3), 213–238.

Kuznetsova A., Brockhoff, P.B., Christensen, R.H.B. (2017). lmerTest Package: Tests in Linear Mixed Effects Models. *Journal of Statistical Software*, *82*(13), 1–26.

Lahiri, A., Wetterlin, A., & Jönsson-Steiner, E. (2005). Lexical specification of tone in North Germanic. *Nordic Journal of Linguistics*, *28*(1), 61–96.

Lee, C. (2007). Does Horse Activate Mother ? Processing Lexical Tone in Form Priming, *50*(1), 101–123.

Malins, J. G., & Joanisse, M. F. (2010). The roles of tonal and segmental information in Mandarin spoken word recognition: An eyetracking study. *Journal of Memory and Language*, *62*(4), 407–420.

Ortega-Llebaria, M., Nemogá, M., & Presson, N. (2017). Long-term experience with a tonal language shapes the perception of intonation in English words: How Chinese-English bilinguals perceive "Rose?" vs. "Rose." *Bilingualism*, *20*(2), 367–383.

Qin, Z., Chien, Y. F., & Tremblay, A. (2017). Processing of word-level stress by Mandarin-speaking second language learners of English. *Applied Psycholinguistics*, *38,* 541–570.

Riad, T. (1998). The origin of Scandinavian tone accents. *Diachronica*, *15*, 63–98.

Riad, T. (2003). Distribution of tonal accents in Scandinavian morphology. In *First International Workshop on Franconian Tone Accent* (pp. 1–12).

Riad, T. (2014). *The phonology of Swedish*. Oxford University Press.

Roll, M., Horne, M., & Lindgren, M. (2010). Word accents and morphology — ERPs of Swedish word processing. *Brain Research*, *1330*, 114–123.

Schremm, A., Novén, M., Horne, M., Söderström, P., van Westen, D., & Roll, M. (2018). Cortical thickness of planum temporale and pars opercularis in native language tone processing. *Brain and Language*, *176*, 42–47.

Schremm, A., Söderström, P., Horne, M., & Roll, M. (2016). Implicit acquisition of tone-suffix connections in L2 learners of Swedish. *The Mental Lexicon*, *11*(1), 55–75.

Sekiguchi, T., & Nakajima, Y. (1999). The Use of Lexical Prosody for Lexical Access of the Japanese Language. *Journal of Psycholinguistic Research*, *28*(4), 439.

Shook, A., & Marian, V. (2016). The influence of native-language tones on lexical access in the second language. *The Journal of the Acoustical Society of America*, *139*(6), 3102–3109.

Singh, L., & Foong, J. (2012). Influences of lexical tone and pitch on word recognition in bilingual infants. *Cognition*, *124*(2), 128–142.

So, C. K., & Best, C. T. (2010). Cross-language perception of non-native tonal contrasts: Effects of native phonological and phonetic influences. *Language and Speech*, *53*(2), 273–293.

Soto-Faraco, S., Sebastián-Gallés, N., & Cutler, A. (2001). Segmental and suprasegmental mismatch in lexical access. *Journal of Memory and Language*, *45*(3), 412–432.

Söderström, P., Horne, M., Frid, J., & Roll, M. (2016). Pre-activation negativity (PrAN) in brain potentials to unfolding words. *Frontiers in Human Neuroscience*, *10*, 1–11.

Söderström, P., Horne, M., & Roll, M. (2017). Stem Tones Pre-activate Suffixes in the Brain. *Journal of Psycholinguistic Research*, *46*(2), 271–280.

van Dommelen, W. A., & Husby, O. (2009). The perception of Norwegian word tones by Chinese and German listeners. *Recent Research in Second Language Phonetics/Phonology: Perception and Production*, (1), 165–174.

Wetterlin, A., Lahiri, A. & Jönsson-Steiner, E. (2007). Tones and loans in the history of Scandinavian. In Tomas Riad & Carlos Gussenhoven (eds.)Tones and Tunes. Volume I: Typological studies in word and sentence prosody. Mouton, 355-377.

Wetterlin, A. (2010). *Tonal accents in Norwegian: Phonology, morphology and lexical specification* (Linguistische Arbeiten, Vol. 535). Walter de Gruyter.

Wiener, S., & Ito, K. (2015). Do syllable-specific tonal probabilities guide lexical access? Evidence from Mandarin, Shanghai and Cantonese speakers. *Language, Cognition and Neuroscience*, *30*(9), 1048–1060.

Yang, J., & Liu, C. (2012). Categorical Perception of Lexical Tone in 6 to 8-year-old Monolingual and Bilingual Children. *International Journal of Asian Language Processing*, *22*(2), 49–62.

Zembrzuski, D., Marecka, M., Otwinowska, A., Zajbt, E., Krzemiński, M., Szewczyk, J., & Wodniecka, Z. (2020). Bilingual children do not transfer stress patterns: Evidence from suprasegmental and segmental analysis of L1 and L2 speech of Polish–English child bilinguals. *International Journal of Bilingualism*, *22*(2), 93–114.