Brain Research 1729 (2020) 146614

Contents lists available at ScienceDirect

Brain Research

journal homepage: www.elsevier.com/locate/brainres

Research report

Different neural mechanisms for rapid acquisition of words with grammatical tone in learners from tonal and non-tonal backgrounds: ERP evidence

Sabine Gosselke Berthelsen[a,*], Merle Horne[a], Yury Shtyrov[b,c], Mikael Roll[a]

[a] Department of Linguistics and Phonetics, Lund University, Sweden
[b] Center of Functionally Integrative Neuroscience, Aarhus University, Denmark
[c] Laboratory of Behavioural Neurodynamics, St. Petersburg State University, Russia

HIGHLIGHTS

- Event-related potentials suggest acquisition of grammatical tone within 20 min.
- Transfer plays a role in how morphosyntactic tone is processed initially.
- Tonal learners draw on native tone-morphosyntax network: early automatic processing.
- Non-tonal learners are not initially able to elicit early processing components.
- Non-tonal learners require consolidation period for mid-latency processing to occur.

ARTICLE INFO

Keywords:
Second language acquisition
Rapid learning
Grammatical tone
Morphosyntax
Transfer
ERP

ABSTRACT

Initial second language acquisition proceeds surprisingly quickly. Foreign words can sometimes be used within minutes after the first exposure. Yet, it is unclear whether such rapid learning also takes place for more complex, multi-layered properties like words with complex morphosyntax and/or tonal features, and whether it is influenced by transfer from the learners' native language. To address these questions, we recorded tonal and non-tonal learners' brain responses while they acquired novel tonal words with grammatical gender and number on two consecutive days. Comparing the novel words to repeated but non-taught pseudoword controls, we found that tonal learners demonstrated a full range of early and late event-related potentials in novel tonal word processing: an early word recognition component (~50 ms), an early left anterior negativity (ELAN), a left anterior negativity (LAN), and a P600. Non-tonal learners exhibited mainly late processing when accessing the meaning of the tonal words: a P600, as well as a LAN after an overnight consolidation. Yet, this group displayed correlations between pitch perception abilities and ELAN, and between acquisition accuracy and LAN, suggesting that certain features may lead to facilitated processing of tonal words in non-tonal learners. Furthermore, the two groups displayed indistinguishable performance at the behavioural level, clearly suggesting that the same learning outcome may be achieved through at least partially different neural mechanisms. Overall, the results suggest that it is possible to rapidly acquire words with grammatical tone and that transfer plays an important role even in very early second language acquisition.

1. Introduction

Learning a new language is often a long and challenging process. It therefore comes as a surprise to many that we are able to deduce simple patterns such as phonotactic rules and indeed even distinguish first words after mere minutes of contact with a foreign language, even if this language is very distant from our own (Gullberg et al., 2010; Gullberg et al., 2012). Our brains perform an impressive task in this endeavour. Even just hearing foreign words as a background stimulus quickly results in changes in our brain activity (Shtyrov et al., 2010; Partanen et al., 2017). This kind of deductive process may be further facilitated when the phonological features of the unknown linguistic input are similar to those of our native language, as opposed to unfamiliar phonology (Kimppa et al., 2015). In the present study, we set out to investigate whether rapid functional changes also materialise in learners' brains during the acquisition of grammatical tone. To this

* Corresponding author at: Center for Languages and Literature, Lund University, Box 201, 22100 Lund, Sweden.
E-mail address: sabine.gosselke_berthelsen@ling.lu.se (S. Gosselke Berthelsen).

https://doi.org/10.1016/j.brainres.2019.146614
Received 22 March 2019; Received in revised form 16 September 2019; Accepted 15 December 2019
Available online 17 December 2019

effect, we use words that express the morphosyntactic features gender and number in tonal and vocalic differences. If rapid acquisition is possible even for complex stimuli with morphosyntactic tone, event-related potential (ERP) components for morphosyntax processing (e.g., left anterior negativity [LAN, Friederici et al., 1993]) should be produced relatively early in the acquisition process. If fast mapping of these novel words and their morphosyntactic tonal content is not possible, we might expect to see a more general, non-morphosyntactic component for language processing (e.g., N400, Kutas and Hillyard, 1980). It is even possible that the acquisition of complex morphosyntactic words is semi-fast, requiring an overnight consolidation for the effects to "grammaticalise", in which case we would observe a LAN on the second day of acquisition but might see an N400 on the first day already. Another feature that might play into the initial acquisition of tonal words is transfer (cf. Section 1.1). If transfer from native (L1) to subsequent languages influences rapid learning, we could expect learners with a tonal L1 to have an advantage over learners with a non-tonal L1. Their native language network for tone processing might allow tonal learners to perceive the tonal information and its familiar function more easily. This could manifest in faster functional changes and in the appearance of early (< 250 ms), pre-attentive ERP components such as the early left anterior negativity (ELAN, Neville et al., 1991) or syntactic mismatch negativity (sMMN, Pulvermüller and Shtyrov, 2003). If, however, transfer does not play a role in rapid learning, we would not expect there to be any differences between learners with a tonal and non-tonal L1. We will consider these issues in more detail.

1.1. Second language acquisition and transfer

There are two categorically different types of language acquisition: first language and second language acquisition (SLA). A speaker's first or native language (L1) is acquired in early childhood, usually starting from birth (Bloomfield, 1933, p. 33) or even earlier. There is evidence from studies with newborns that suggests that some basic prosodic processing, e.g., recognition of prominent melodic patterns, is possible even before birth (Mampe et al., 2009). Speakers can have two or more native languages, which can be acquired simultaneously or successively (Tsimpli, 2014). Yet, there is a cut-off point for native language development possibly as early as age 4 (Meisel, 2009). A language acquired after the native language cut-off point is called second language (L2). Second language learners can be children (early SLA) and adults (late SLA). In recent years, there has been a growing interest in the very initial stages of (adult) SLA (cf. Gullberg and Indefrey, 2010). A variety of studies have been able to show behavioural (e.g., Gullberg et al., 2010) and electrophysiological evidence (e.g., Partanen et al., 2017) for successful learning within very short periods of time, i.e., 20 min or less. However, it has also been suggested that learned linguistic material needs to be consolidated during sleep (overnight consolidation) in order to fully be processed like real language (Davis et al., 2009).

An important parameter in second language acquisition is transfer. When learning any non-native language, learners already possess a pre-existing language system which is based on their native language and other languages that they might previously have learned (Rast, 2010). New languages are built upon the architecture of this pre-existing system, which in turn influences the way subsequent languages are acquired: a concept that is referred to as cross-linguistic influence or transfer (Lado, 1957). Considering, for instance, the acquisition of number inflections on nouns (Charters et al., 2012; Luk and Shirai, 2009), it has been shown that when nouns are not inflected for number in a learner's L1, the learner will struggle with this feature in his/her L2. This is an example of negative transfer. Importantly, transfer can also be positive: e.g., learners that do have nominal number in their L1 will find it easy to produce this in their L2 (Luk and Shirai, 2009). Transfer effects have been observed in virtually all aspects of language: pragmatics (Bou-Franch, 2013; Kasper, 1992), semantics (Ghazi-Said and Ansaldo, 2017), syntax (Hawkins and Lozano, 2006; Rothman and Cabrelli Amaro, 2010), morphology (Hawkins and Lozano, 2006; Ramirez et al., 2013), and phonology (Hawkins and Lozano, 2006; Simon, 2010), including prosody (Jun and Oh, 2000; Lee and Matthews, 2015). In the initial stages of the acquisition of a second language, it is conceivable that transfer plays a particularly important role (Rast, 2010; Shih and Lu, 2010). Without any knowledge of the foreign language, one has to rely on one's own language system in order to unravel the structure of the unfamiliar input. Thus, we believe transfer may play an important role in the present study where learners are introduced to a new language with word-level tones.

1.2. Tone languages and the acquisition of tonal features

1.2.1. Types of tone

Linguistic tone is a prominent feature in the world's languages with 40–70% of all languages being tonal (Maddieson, 2013; Yip, 2002). Tones are most commonly added to words to make lexical distinctions (i.e., lexical tone, e.g., Chinese: Yip, 1980, Vietnamese: Alves, 1995) or grammatical distinctions (i.e., grammatical tone, e.g., Hausa: Crysmann, 2015, Dogon: McPherson, 2014, or Somali: Banti, 1989; Le Gac, 2003). There are also languages where word-level tones (so-called 'word accents') are neither intrinsically lexical nor grammatical in themselves but rather associated with different kinds of morphosyntactic information. An example of this type of language is Swedish (cf. Bruce, 1977; Riad, 1998, but also Riad, 1998; Rischel, 1963 for Norwegian), where tones on stems are for instance associated with certain singular–plural distinctions or present–past tense distinctions specified by suffixes. On a nominal word stem (e.g. *hatt*, hat), a low tone (i.e. accent 1 → *hatt*$_1$) will cue the listener to expect an upcoming definite singular suffix (*hatt*$_1$*en*, hat-the) rather than a plural suffix (**hatt*$_1$*ar,* *hats) or a productive compound (**hatt*$_1$*hylla,* *hat rack). Plural *-ar* and compounds instead have to be preceded by a high tone (i.e. accent 2 → *hatt*$_2$*ar*, hats; *hatt*$_2$*hylla,* hat rack) on the nominal word stem. As such, the tone-suffix relation in a word accent language like Swedish entails that morphosyntactic information can be pre-activated by word-stem tones in speech perception (e.g., Söderström et al., 2016; Roll et al., 2015).

1.2.2. The acquisition of L2 tone

The acquisition of tone poses a great challenge for L2 learners, and native-like proficiency appears hard to reach (Yang and Chan, 2010). In light of transfer and a conceivably fine-tuned tone perception and production system in speakers with a tonal L1, it would seem natural that such learners should have an advantage in the acquisition of a tonal L2. However, this is not necessarily supported by the data. While some studies have found an advantage for tonal L1s in the perception (Francis et al., 2008; Hallé et al., 2004; Lee et al., 1996; van Dommelen and Husby, 2007; Wayland and Li, 2008) and production (Zetterholm and Tronnier, 2012) of L2 tone, others have not found it for either perception (Chen et al., 2015; Francis et al., 2008; Hao, 2012; Lee et al., 1996) or production (Tronnier and Zetterholm, 2015). It seems that three main factors influence whether or not transfer can result in a positive effect on tone perception and acquisition: tonal complexity, tonal similarly, and task difficulty. If the native tone system is more complex than the target system, there is a high chance that the learners will show transfer-based advantage compared to learners with less complex native tone systems or those without a native tone (cf. Lee et al., 1996). Similarly, for target tones that are perceptually or functionally similar to learners' native language tones, transfer seems to be possible (e.g., Braun and Johnson, 2011; So and Best, 2010). These two internal factors, collectively, speak to the tuning and nuancing of the perceptive and functional tone processing system. The third factor, task difficulty, is external to language, although it is still related to the internal factors. Tone perception or acquisition tasks that are fairly simple (e.g., discrimination of two distinct tones) do not require advanced perceptive processing and, therefore, will often not show behavioural differences between learners of tonal and

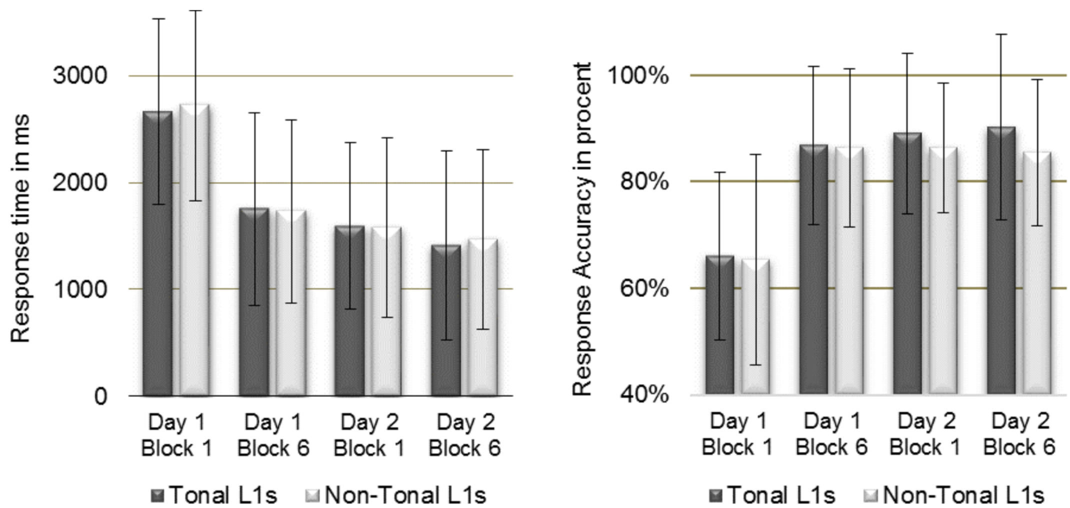

Fig. 1. Mean response time (left) and response accuracy (right) at the beginning and end of both days by group (bars represent standard deviations).

non-tonal languages (cf. Chen et al., 2015; Lee and Matthews, 2015), presumably due to ceiling effects. It stands to reason, however, that the processing mechanisms might still be different even though the behavioural data show similar performance (e.g., Wang et al., 2001). More complex tasks (for instance, discrimination of tones within a tone continuum, Hallé et al., 2004, or discrimination with disrupting task, Lee et al., 1996), on the other hand, have been observed to lead to the emergence of visible transfer effects even in behavioural measures. Importantly, regardless of potential transfer effects, both the perception (Francis et al., 2008; Hed et al., 2019; Schremm et al., 2017; Shen and Froud, 2016; Shih and Lu, 2010; van Dommelen and Husby, 2007; Wang et al., 1999; Wayland and Li, 2008) and the production of L2 tone (Wang et al., 2003) can be improved by training. Training has even been seen to lead to retainment and generalizable categorical knowledge (Wang et al., 1999). Note that all of the cited studies were carried out on L2 languages with either lexical tones or word accents. There is to our knowledge only a single study on the acquisition of a language with grammatical tone (Yoruba) by Orie (2006). This study found that tones which typically carry morphosyntactic function (i.e., mid tones) are those which are most difficult to acquire (3% accuracy versus 73% and 88% for high and low tones, which are predominantly used as lexical tones in Yoruba). However, this might also at least partly be due to the fact that mid tones had no direct equivalent in the learners' L1 prosodic system (i.e., English).

Having a strong connection between tones and morphosyntax, we propose that native speakers of a word accent language may have an advantage in the acquisition of a language with grammatical tone. Although different information is conveyed by word accents (cues to upcoming grammatical suffixes) and grammatical tones (grammar), there are strong ties to morphosyntax in both cases. The neural subsystems used in the processing of morphosyntax-related tones may give L1s from a word accent language (here, Swedes; henceforth Tonal L1s) an advantage over learners from a language without tone (here, Germans[1], henceforth Non-Tonal L1s) in the perception of L2 tones and the association of tonal L2 words and their morphosyntactic meaning. If the corresponding neural sub-systems can be taken advantage of even in the earliest phases of acquisition, we should see different patterns in the behavioural and electrophysiological responses between Tonal L1s and Non-Tonal L1s.

1.3. ERP correlates of morphosyntactic processing in L1 and L2

Our brain processes language in a very structured way, and we are able to observe typical language processing patterns with the help of electroencephalographic recordings (EEG) and event-related brain potential analyses. ERPs for language are most prominently observable in contrasts between legal strings and illegal ones (where linguistic processing fails).

1.3.1. Early ERP components

At around 30–80 ms post-stimulus identification point, where the auditory input has just reached the language processing areas and language processing becomes possible, a relatively novel ERP effect is found (Shtyrov and Lenzen, 2017). This early ERP (falling within the same range as P1/P50 response) indexes the first stage of automatic lexical access (MacGregor et al., 2012; Shtyrov and Lenzen, 2017), where the listener can for instance distinguish between words and pseudowords (Sainz and Lazaro, 2009). A modulation of this early effect has also been observed in the context of learning, where its amplitude changes with repeated exposure to novel words (Kimppa et al., 2015, Kimppa et al., 2016). Similarly, this effect emerges for trained novel words compared to untrained novel words (Leminen et al., 2016).

Another early processing component is the early left anterior negativity (ELAN), which has often been observed in studies with morphosyntactically strongly dispreferred phrase continuations (e.g., Friederici et al., 1993; Hahne and Friederici, 1999; Hahne and Friederici, 2002; Neville et al., 1991) which may be perceived as incorrect.[2] The ELAN is traditionally considered an indicator of word category violation (Friederici, 2002). ELANs are rarely found in studies with second language learners (L2s) (e.g., Hahne, 2001; Hahne and Friederici, 2001). In fact, even learners who acquired a language very early in life do not produce early left anterior negativities that resemble those of native speakers (Weber-Fox and Neville, 1996). Yet, learners with very high proficiency in a language - artificial (Friederici et al., 2002) or natural (Hanna et al., 2016) - have been reported to produce ELANs.

1.3.2. Late ERP components in morphosyntax processing

Besides the early, more automatic components, language processing also elicits later ERPs. For morphosyntactic discrepancies, for instance, L1s will elicit a left anterior negativity (LAN) at around 300 – 500 ms after the emergence of the inadmissible linguistic material (Molinaro

[1] German was chosen as the non-tonal language because of its typological and phonological similarity to Swedish. The phoneme inventory in German and Swedish is very similar (see supplementary material for full phoneme inventory based on Wiese (1996) and Riad (2014)). German and Swedish intonation are also relatively similar; however, German does not have word-level tones like Swedish (Gårding, 1998; Gibbon, 1998).

[2] Refer to Steinhauer and Drury (2012) for a discussion about the grammaticality of sentences used in ELAN studies.

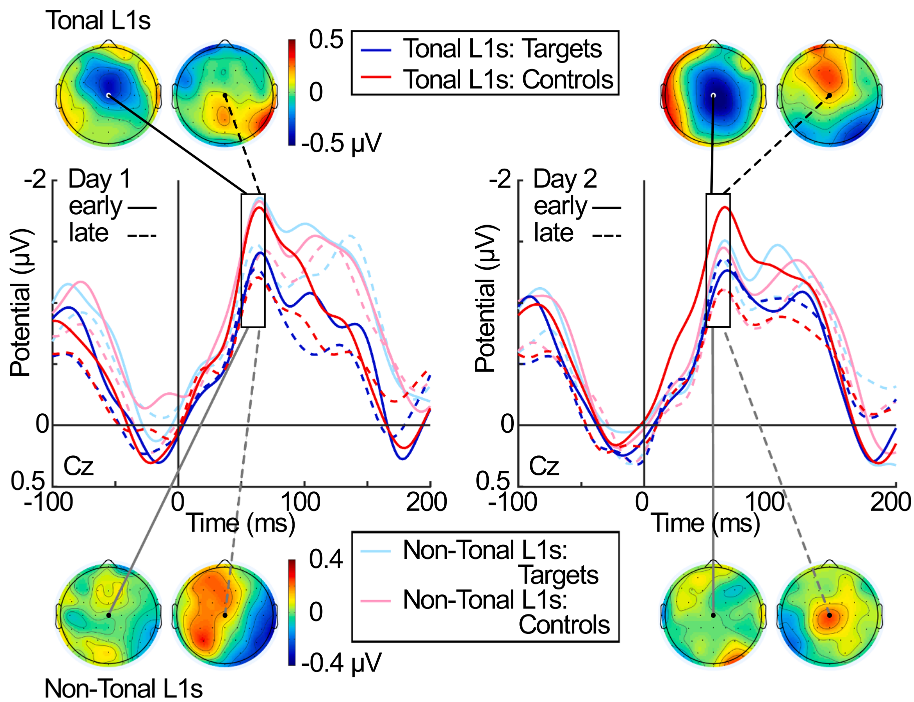

Fig. 2. Middle: ERPs from both groups at central electrode Cz for early and late trials on Day 1 (left) and Day 2 (right). Tonal L1s in dark colours, Non-Tonal L1s in pale colours. Early trials in continuous lines, late trials in dotted lines. Above and below: Topographical distributions for the effects (controls-targets) at both time points on Day 1 (left) and Day 2 (right) for Tonal L1s (above) and Non-Tonal L1s (below).

et al., 2014). In second language learners, the LAN is only rarely observed, yet highly proficient learners have been seen to generate this effect (Steinhauer et al., 2009). Less proficient learners have rather been seen to elicit an N400 in response to morphosyntactic errors (e.g., McLaughlin et al., 2004). The N400 normally serves as an indicator of failed semantic rather than morphosyntactic processing (cf. e.g., Kutas and Hillyard, 1980; Kutas and Federmeier, 2014).

Following the mid-latency negativity, a late positive component (P600) emerges, in particular during syntactic processing (Osterhout and Holcomb, 1992). This late deflection has a centroparietal distribution and is most commonly considered an indicator of failed morphosyntactic integration (Kuperberg, 2007). As such, it is believed to signal the need for revision or repair (Hagoort, 2003; Kaan and Swaab, 2003). Second language learners are frequently seen to produce P600s (Gillon Dowens et al., 2009; Sneed German et al., 2015).

The ELAN, LAN, and P600 are best observed in cases of failed processing, although these components appear to be general indicators of different kinds of language processing which can emerge (albeit to a smaller extent, cf. e.g., Krott and Lebib, 2013) even without morphosyntactic violations.[3] We therefore expect a P600, LAN, and ELAN for morphosyntactically charged artificial target words compared to morphosyntactically and semantically empty control words.

It is believed that a P600, LAN, and ELAN cannot be elicited before the morphosyntactic features of the target language have become grammaticalised in the learner's brain (Steinhauer, 2014). In natural second language acquisition, it can take many years for learners to reach sufficiently high proficiency in morphosyntax processing to produce these ERPs. Late learners who were reported to produce a LAN and P600 in Gillon Dowens et al. (2009), for instance, had an average of 22 years of immersion in the target language environment. By means of using strongly constrained artificial grammars, however, the advance to high proficiency can be tremendously accelerated. In Silva, Folia, Hagoort & Petersson (2017), for example, a centroparietal P600 was found after a mere 4 h of grammar training. Participants in this study, however, did not produce a LAN, which seems to require an even higher level of proficiency (Steinhauer et al., 2009) and/or a different, more combinatorial morphosyntax processing strategy (Rodriguez-Fornells et al., 2001). A different study with an artificial grammar, where exact length of training was not reported but clearly exceeded 4 h, found both an ELAN, a late negativity and a P600 (Friederici et al., 2002). Steinhauer et al. (2009) argue for the importance of proficiency (rather than length of acquisition) and propose a progression of morphosyntactic processing in second language learners, from no linguistic ERP effects in novices via N400 for low proficiency learners to P600 and eventually LAN and ELAN for high proficiency L2s. Different forms of processing at different proficiency levels (whole word processing in mid-proficiency learners versus rule-based processing in advanced learners) may explain the shift from late positive components to a biphasic activation pattern.

In the present study, we investigate whether it is possible to rapidly reach high proficiency when acquiring a small set of novel tonal words in two 2-hour-long word-picture-association sessions on consecutive days. We monitor behavioural changes as well as language ERP components that could be expected to occur within minutes of exposure and after an overnight consolidation stage, and assess whether there are any changes for tonal words. We test a group of learners who have a functionally similar but more complex tone system in their native language in order to examine which processing patterns emerge for learners who should be expected to rely on transfer mechanisms for processing word tones. This allows us to study whether transfer based on tone-morphology processing subsystems is possible in this early acquisition phase. Additionally, we test a separate group of learners whose L1 lacks word tones in order to see how an assumed no-transfer group initially deals with words with morphosyntactic tone. We specifically investigate what type of processing learners, whose native language does not have tones at the word level, will refer to and how this may be reflected in their behavioural results. We expect that the Non-Tonal L1 group will compensate with, e.g., general pitch perception abilities due to not having a network for processing tone at the word level and that they might be relatively heterogeneous regarding this possibility. We therefore also included a background experiment, which specifically targeted the participants' pitch perception abilities. Given the potential source of heterogeneity in the Non-Tonal L1 group,

[3] See Pulvermüller and Shtyrov (2003) for a discussion on grammatically correct strings in (at least early) ERPs.

we tested the two groups separately, compared their results with the help of t-tests, and further disentangled the Non-Tonal group's results with the help of correlation analyses.

2. Results

2.1. Behaviour

In the Tonal L1 group, RTs in question trials showed a Day*Block interaction, $F(1,22) = 9.06$, $p = .006$. The Tonal L1 participants' response times became significantly faster from the beginning ($M = 2661$ ms, $SD = 886$) to the end ($M = 1751$ ms, $SD = 856$) of the first day, $F(1,22) = 22.03$, $p < .001$. For Response Accuracy, there was also a Day*Block interaction, $F(1,22) = 35.83$, $p < .001$. The Tonal L1 participants' responses improved significantly in accuracy from early ($M = 65.98\%$, $SD = 19.82$) to late ($M = 86.85\%$, $SD = 14.92$) trials on day 1, $F(1,22) = 42.46$, $p < .001$. See Fig. 1.

In the Non-Tonal L1 group, RTs in question trials similarly revealed a Day*Block interaction, $F(1,22) = 11.88$, $p = .002$. The Non-Tonal L1 participants' response times sped up significantly from the beginning of the first day ($M = 2719$ ms, $SD = 867$) to the end ($M = 1729$ ms, $SD = 902$) of the first day, $F(1,22) = 23.71$, $p < .001$. Likewise, there was a Day*Block interaction, $F(1,22) = 66.41$, $p < .001$, for the behavioural measure Response Accuracy. Thus, the Non-Tonal L1 group's accuracy for question trials significantly improved between early ($M = 65.36\%$, $SD = 15.74$) and late ($M = 86.38\%$, $SD = 14.89$) trials on day 1, $F(1,22) = 82.47$, $p < .001$. See Fig. 1.

The acceleration of Response Time between the beginning and end of day 1 did not differ significantly between Tonal L1 ($M = -909$ ms, $SD = 929$) and Non-Tonal L1 ($M = -989$ ms, $SD = 974$) participants; $t(44) = 0.28$, $p = .777$. Similarly, there was no significant difference in the improvement of Response Accuracy from early day 1 to late day 1 between Tonal L1 ($M = +20.87\%$, $SD = 15.36$) and Non-Tonal L1 ($M = +21.02\%$, $SD = 11.10$) learners; $t(44) = 0.38$, $p = .970$.

2.2. Electrophysiology

For the 50–70 ms time window, there was a Word Type*Block*Laterality interaction, $F(2,44) = 5.76$, $p = .007$, for the Tonal L1 group. Subsequent separate rmANOVAs for early and late trials revealed a Word Type*Laterality interaction, $F(2,44) = 5.69$, $p = .014$, for early trials. This was based on a negativity for control words at mid-lateral electrodes, $F(1,22) = 9.20$, $p = .006$ (Fig. 2). There were no significant main for effect or interactions with Word Type for the Non-Tonal L1 group in this time window. The Tonal L1 ($M = 0.01$ µV, $SD = 0.25$) and Non-Tonal L1 ($M = -0.23$ µV, $SD = 0.36$) participants differed significantly in their potentials for controls minus targets for early trials at mid electrode sites, $t(44) = 2.66$, $p = .011$.

For the 145–165 ms time window, the Tonal L1 group showed a Word Type*Laterality interaction, $F(2,44) = 3.97$, $p = .027$, which was due to a negativity at left-lateral electrodes, $F(1,22) = 8.19$, $p = .009$, and a positivity at right-lateral electrodes, $F(1,22) = 4.88$, $p = .038$, for targets compared to controls, cf. Fig. 3. There were no significant main effect for or interactions with Word Type in the Non-Tonal L1 group for this peak. Yet, the Tonal L1 ($M = -0.09$ µV, $SD = 0.16$) and the Non-Tonal L1 ($M = -0.09$ µV, $SD = 0.24$) participants did not differ significantly in their amplitude for targets minus controls at left lateralities, $t(44) = 0.05$, $p = .962$. A subsequent correlation analysis revealed a significant negative correlation of the Non-Tonal L1 participants' amplitude at left-lateral electrodes with Accuracy in Non-Linguistic Pitch Distinction, $r = -0.547$, $p = .012$ (Bonferroni-corrected), cf. Fig. 4. Mean Pitch Distinction Accuracy in the Non-Tonal L1 group was 93.75% with a range from 79.69% to 100%; $SD = 6.20$.

For the 460–560 ms time window, there was a Word Type*Posteriority interaction, $F(2,44) = 19.08$, $p < .001$, for the Tonal L1 group, which stemmed from an anterior negativity, $F(1,22) = 15.70$, $p < .001$, and a posterior positivity, $F(1,22) = 22.74$, $p < .001$, for targets compared to control words. The Non-Tonal L1 group showed a Word Type*Day*Posteriority interaction, $F(2,44) = 4.20$, $p = .041$, in this time window. Separate rmANOVAs for day 1 and day 2 subsequently revealed a Word Type*Posteriority interaction for day 2 which was based on an anterior negativity, $F(1,22) = 7.24$, $p = .013$, and a posterior positivity, $F(1,22) = 12.11$, $p = .003$, for targets compared to controls, cf. Fig. 5. However, there was no significant difference in the ERP amplitudes for target minus control words at left lateral electrodes on day 1, $t(44) = 1.14$, $p = .259$, between the Tonal L1 ($M = -0.32$ µV, $SD = 0.61$) and Non-Tonal L1 ($M = -0.12$ µV, $SD = 0.55$) participants. A subsequent correlation analysis for the Non-Tonal L1 participants' mean amplitudes for targets minus controls at anterior sites on day 1 revealed a negative correlation with Response Accuracy, $r = -0.469$, $p = .048$ (Bonferroni-corrected), cf. Fig. 4. Mean Response Accuracy in the Non-Tonal L1 group was 79.15% but had a relatively large range from 52.45% to 97.69%; $SD = 12.86$.

Since the effect at 460–560 ms in the Non-Tonal L1 group emerged on day 2, we conducted an overnight consolidation analysis of the ERP data in this time window. There were no significant effects for Consolidation in the Tonal L1 group. In the Non-Tonal L1 group, there was a Consolidation*Posteriority interaction, $F(2,44) = 4.44$, $p = .041$, which was driven by an overall reduced negativity after consolidation at anterior electrodes, $F(1,22) = 5.29$, $p = .031$, cf. Fig. 6. The difference in amplitudes at anterior electrodes between Tonal L1s ($M = -0.07$ µV, $SD = 1.14$) and Non-Tonal L1s ($M = 0.55$ µV, $SD = 1.14$) falls just short of significance, $t(44) = 1.83$, $p = .075$. There were no significant correlations between amplitude at anterior electrodes with behavioural factors in either group. We believe the lack of between-group differences to be based on high intra-group variation.

For the 670–770 ms time window, the Tonal L1 group showed a Word Type*Posteriority interaction, $F(2,44) = 15.00$, $p = .001$, which was based on an anterior negativity, $F(1,22) = 10.06$, $p = .004$, and a central, $F(1,22) = 5.09$, $p = .034$, and posterior positivity, $F(1,22) = 21.10$, $p < .001$, for targets compared to controls, cf. Fig. 7. The same effect was seen in the Non-Tonal L1 group, where there was a Word Type*Posteriority interaction, $F(2,44) = 9.67$, $p = .004$, which broke down into an anterior negativity, $F(1,22) = 6.38$, $p = .019$, and a central, $F(1,22) = 9.88$, $p = .005$, and posterior positivity, $F(1,22) = 13.13$, $p < .00$, for targets compared to controls, cf. Fig. 7. T-tests show no significant differences between groups, neither at anterior, $t(44) = 0.47$, $p = .640$, nor at central, $t(44) = 0.26$, $p = .795$, or posterior, $t(44) = -0.45$, $p = .657$, sites.

3. Discussion

Using electroencephalographic recordings, the present study set out to investigate whether speakers from both tonal and non-tonal backgrounds can rapidly acquire non-native words with morphosyntactic tonal features. We used words from an artificial language that expressed gender and number through vocalic and tonal contrasts. The words were taught in a sound-picture association task and compared to repeated, non-meaningful control stimuli. Results showed that speakers with a tonal background could make use of early, automatic neural processing to assess the novel words and were significantly faster to show differences in later processing as well. Speakers with a non-tonal background relied mainly on late processing components to access the meaning of the novel words. Interestingly, despite the obviously different processes involved in the acquisition, there were no between-group differences in the behavioural performance. In the following paragraphs, we will first summarise the early and late components that we found in the Tonal L1 group. Afterwards, we will discuss how the Non-Tonal L1 results compare.

3.1. Early, automatic components

3.1.1. Word recognition ERP for novel words and novel pseudowords

The first ERP effect we found between target words and control

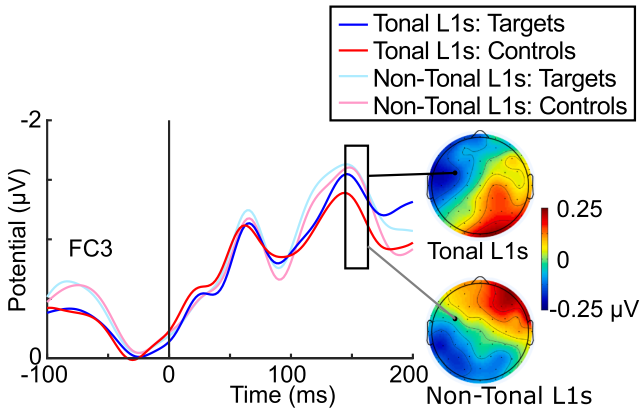

Fig. 3. Left: ERPs at left-frontal electrode FC3 for the two groups for all trials: Tonal L1s in dark colours, Non-Tonal L1s in pale colours. Right: Topographical distributions of target-control effects at ELAN latency (145–165 ms) for all trials in the Tonal L1 (above) and Non-Tonal L1 (below) group.

words was an effect at around 50–60 ms that is associated with word recognition. Previous studies have found clear differences in the ERP for known words and pseudowords in this time window (Kimppa et al., 2015; MacGregor et al., 2012; Shtyrov and Lenzen, 2017); these differences can even out towards the end of the recording session, which is believed to be a sign of memory-trace build up (Kimppa et al., 2015). The present study found a central negativity, session-initially, for novel, tonal pseudowords compared to novel, tonal words. It appears as though the tonal features allowed the native tonal listener to instantaneously single out the pseudowords as filler items that they would not need to allocate attention to. The novel target words, on the other hand, were harder to distinguish as they had a novel semantic and morphosyntactic load that the learner needed to connect to the tonal and segmental features. The effect at this early latency was *only* observed in learners with a *tonal L1* and disappeared towards the end of the session. Interestingly, the effect re-appeared session-initially on day 2, suggesting that control words were not consolidated.

3.1.2. ELAN for novel words

The second early ERP component that we found for target words compared to controls was an early left anterior negativity. Given the lack of violations in the input, the emergence of the ELAN may seem unexpected at first. Yet, it fits in perfectly when we consider the distinguishing characteristic of processing at this latency: the automatic activation of a morpheme's syntactic properties (cf. Bakker et al., 2013).

Largely depending on stimulus design, a number of different negative components emerge at around 150 ms post stimulus for morphosyntactic processing: the ELAN, the syntactic mismatch negativity (sMMN), and the pre-activation negativity (PrAN) (Roll et al., 2017; Söderström et al., 2016). The ELAN is ordinarily observed with sentential stimuli and its amplitude varies as a function of syntactic pre-activation. In sentences where the studied morpheme follows the most likely and thus pre-activated phrase structure, the ELAN amplitude is reduced as compared to sentences where a morpheme from a non-pre-activated word category appears (e.g., Friederici, 2002; Neville et al., 1991). The ELAN is not readily replicable but has been observed in different languages (e.g., English: Neville et al., 1991, German: Friederici, 2002, French: Isel et al., 2007) and for different phrase structure types (e.g., discontinued prepositional phrase, Neville et al., 1991, discontinued verb phrase, Friederici, 2002). The sMMN, much like the ELAN, is a component that shows reduced neural activity in cases where morphosyntactic structures are pre-activated. Following a classical MMN design where a deviant stimulus interrupts repetitions of a standard stimulus, listeners exhibit a decreased sMMN for grammatically congruent deviants as compared to ungrammatical deviants. This rather constant effect has been observed for a number of different languages (e.g., English: Pulvermüller and Shtyrov, 2003, German: Pulvermüller and Assadollahi, 2007, Finnish: Shtyrov et al., 2007) and agreement types (e.g., pronoun-verb agreement, Pulvermüller and Shtyrov, 2003, determiner-verb agreement, Hasting et al., 2007, and plural inflections, Bakker et al., 2013). Reduction of both the ELAN and the sMMN are believed to show the eased activation of a grammatically valid string due to successful priming through pre-activation. The PrAN, finally, differs from the two above-mentioned components in so far as its amplitude is not indicative of the pre-activated status of the present morpheme. In fact, the studied morphemes in a PrAN paradigm are designed to be equally strongly pre-activated. Instead, the stimuli for which a PrAN is elicited are manipulated so that they themselves – upon the activation of their morphosyntactic properties – pre-activate subsequent morphemes to different degrees. The PrAN's amplitude varies as a function of the number of possible continuations, i.e., based on the morpheme's predictive strength. What all three components of morphosyntactic processing at this latency appear to share is that they strongly depend on access to the input's morphosyntactic content. We believe their interaction and respective roles to be the following: whereas the ELAN and the sMMN are indicative of the processing effort that is required to activate a morpheme's morphosyntactic content based on whether or not it was pre-activated, the PrAN subsequently demonstrates to what extent the activation of said morphosyntactic

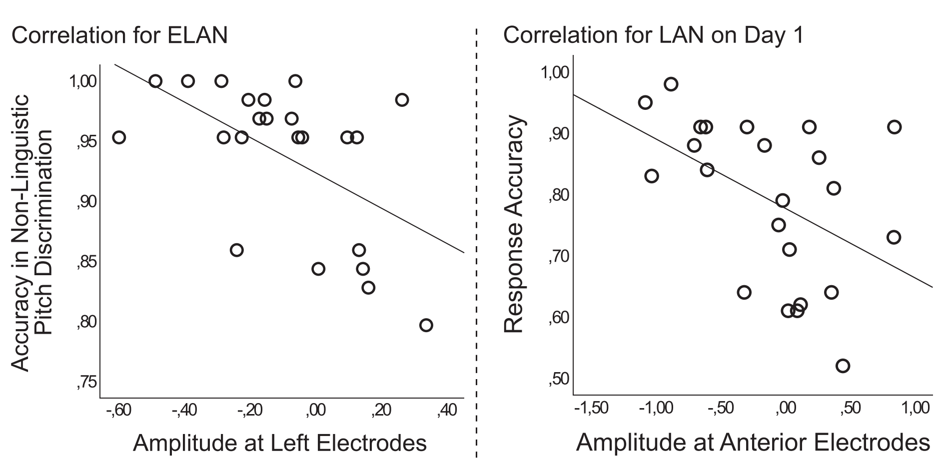

Fig. 4. Left: Scatterplot for the correlation of negativity at left lateral electrodes for targets-controls and Accuracy in Non-Linguistic Pitch Distinction in the Non-Tonal L1 group. Right: Scatterplot for the correlation of negativity at anterior electrodes for targets-controls on day 1 and Response Accuracy in the Non-Tonal L1 group.

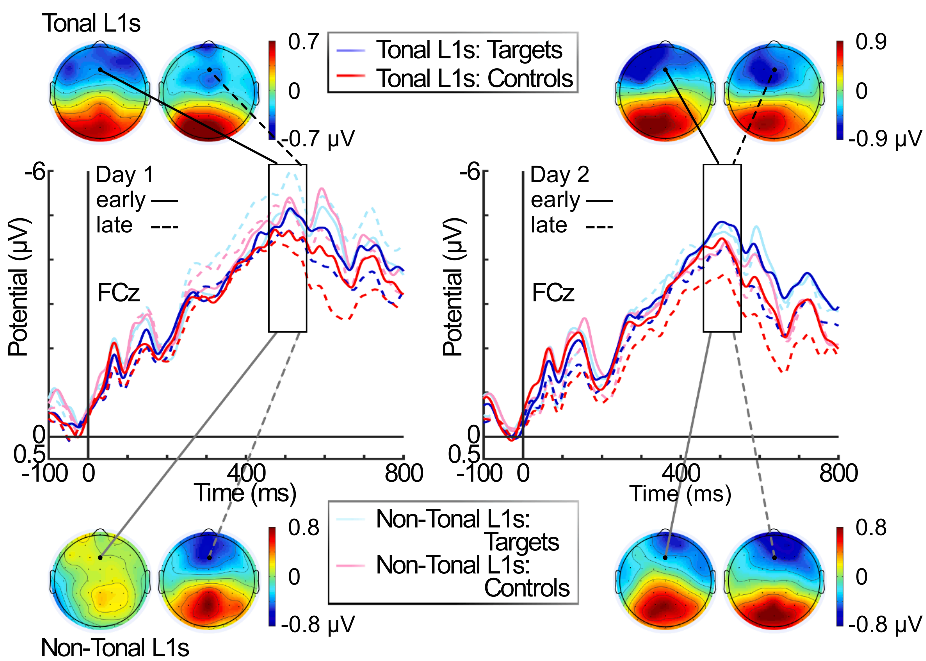

Fig. 5. Middle: ERPs from both groups at mid-frontal electrode FCz for early and late trials on Day 1 (left) and Day 2 (right). Tonal L1s in dark colours, Non-Tonal L1s in pale colours. Early trials in continuous lines, late trials in dotted lines. Above and below: Topographical distributions for the effects at both time points on Day 1 (left) and Day 2 (right) for Tonal L1s (above) and Non-Tonal L1s (below).

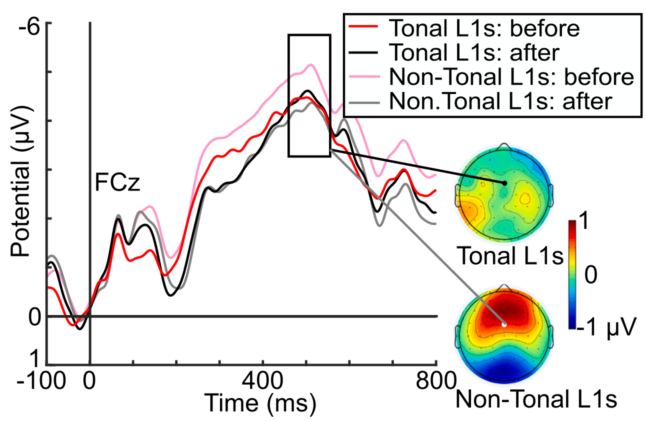

Fig. 6. Left: ERPs at anterior electrode FCz for the two groups for all trials: Tonal L1s in dark colours, Non-Tonal L1s in pale colours. Right: Topographical distributions of consolidation effects (after-before consolidation) at 460–560 ms for all trials in the Tonal L1 (above) and Non-Tonal L1 (below) group.

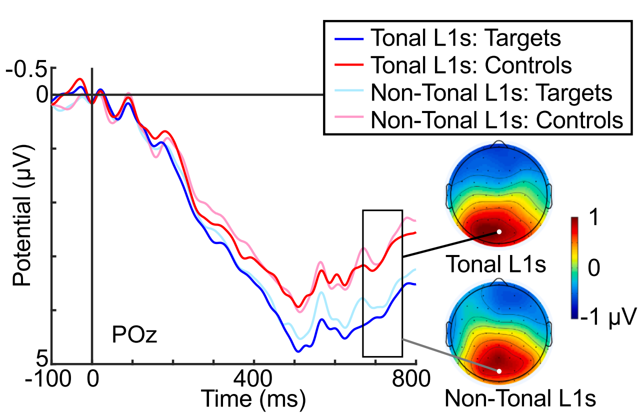

Fig. 7. Left: ERPs at posterior electrode POz for the two groups for all trials: Tonal L1s in dark colours, Non-Tonal L1s in pale colours. Right: Topographical distributions of target-control effects at P600 latency (670–770 ms) for all trials in the Tonal L1 (above) and Non-Tonal L1 (below) group.

properties immediately makes predictions about possible continuations. Based on the above discussion of the morphosyntactic processing at around 150 ms post stimulus, we consider the ELAN to be an indicator of morphosyntactic activation that is readily influenced by pre-activation. Violations of pre-activated structures will hamper morphosyntactic activation and, therefore, elicit a larger ELAN response than canonical sentences. However, being an indicator of morphosyntactic activation, as we suggest, an ELAN should also appear in non-violating contexts, i.e., when comparing contexts that elicit morphosyntactic activations to contexts that do not.

The stimuli in the present study are equally strongly pre-activated, as the context preceding the disambiguation point (to which ERPs are time-locked) is equally frequently followed by all possible continuations. Furthermore, the stimuli have no difference in predictive strength, as only one continuation is possible after the word has been disambiguated. The fact that we find a larger ELAN for morphosyntactically loaded target words compared to morphosyntactically empty controls in the current data can solely be explained by the fact that morphosyntactic activation takes place in one condition but not the other. Larger neural activity is required for accessing the number and gender properties of the target words while lack of number and gender results in a reduced ELAN amplitude for control words.

Interestingly, already within the first 20 minutes of aquisition, we found an ELAN which did not differ significantly from the ELANs produced at the following three time points that were analysed for learning effects. This suggests that the Tonal L1s were capable of deducing morphosyntactic features already within just five repetitions of all stimuli. This is a somewhat surprising finding, since components for morphosyntactic processing are not even commonly found in highly proficient L2 learners (e.g., Hahne and Friederici, 2001; Weber-Fox and Neville, 1996) or early bilinguals (Weber-Fox and Neville, 1996). However, ELAN-eliciting paradigms used in previous studies with L2 learners depend on whether or not morphosyntactic activation was facilitated by pre-activation. Predictive processing, which would lead to pre-activation, is not frequently observed in L2 learners (Kaan, 2014). In the present study, the ELAN was elicited for morphosyntactic activation, unconditioned by pre-activation, which is something the Tonal L1 speakers, owing to L1-L2 transfer, could apparently achieve in the restricted context of our paradigm. Based on the relative simplicity of the stimuli, the Tonal L1 learners in our study could perceive the linguistic importance of the novel tones, which allowed them to quickly access and activate the words' morphosyntactic features. This type of early automatic morphosyntactic processing for novel morphosyntactic information, when compared to the processing of non-meaningful pseudoword controls, resulted in an ELAN.

3.2. Late components

3.2.1. LAN for novel tonal words

The third component analysed in the present study was a left anterior negativity for novel tonal words. The LAN is a well-known component for morphosyntactic processing. As mentioned in the introduction, the LAN is only found in second language learners with very high levels of proficiency. Yet, learners can very rapidly reach high proficiency when acquiring small-scale artificial grammars. In the present study, we found a LAN for the processing of morphosyntactic information. The effect presumably indexed the rule-based integration of morphosyntax with the words' semantic information. There were no interactions with temporal factors which suggests that the Tonal L1s were capable of morphosyntactic processing using tonal information after just 20 minutes of word-picture learning. Interestingly, their response accuracy (66%) showed that they did not perform much above chance level behaviourally. For a comparable result in a tone-morphology training experiment, see Hed et al. (2019). Simple rule-morphology training yielded a LAN in learners resembling that for pseudowords in native speakers, which suggested that the learners acquired the tone-morphology association before they had grasped the semantics. This leads us to believe that the simplicity of the stimuli in the present study and the Tonal L1s native tone-morpheme matching system allowed them to make an automatic tone-morphosyntax connection even before they could consciously perceive the differences (cf. McLaughlin et al., 2004; Tokowicz and MacWhinney, 2005).

3.2.2. P600 for novel tonal words

The final component we analysed was a P600. The P600 is a marker of integration of the current word into the current context model and revision or repair. In the present study, we found a P600 for our multimorphemic target words outside of sentential context and in the absence of violation. We believe the P600 in this case to be a marker of unification of the words' morphosyntactic, semantic and also segmental and prosodic properties (e.g., whole-word melody contour) and formation of a holistic memory trace. The P600 was found for all time points in the experiment and for both speaker groups alike. The emergence of both early but importantly also later, morphosyntactically charged components in the study shows that rapid acquisition indeed is possible even for multi-layered linguistic material such as words with grammatical tones. The appearance of a LAN rather than an N400 strongly suggests that the grammatical, and not only the semantic content of the stimuli, is acquired within a few minutes of learning in the Tonal L1 group. This allows us to argue that rapid acquisition before overnight consolidation is possible even for linguistically complex, combinatorial stimuli when learners can profit from positive L1-L2 transfer.

3.3. The role of L1 to L2 transfer in the acquisition of novel tonal words

Having described in detail how native speakers from a language with morphosyntax-related tone respond to novel morphosyntactically charged tonal words, we will now consider how the lack of positive L1-L2 transfer affects acquisition and processing of said words in Non-Tonal L1s: Unlike the Tonal L1s, Non-Tonal L1s did not elicit a differential effect for controls and targets at word recognition latency (~50 ms). It stands to reason that their auditory language processing system was not tuned to picking up the tonal cues to word status and that the vocalic cues on their own were not strong enough to allow for rapid word status recognition and functional allocation of attentional resources.

Similarly, the Non-Tonal L1s elicited no significant effect in the ELAN time window. Yet, there was a negative correlation of negativity at ELAN latency/location and the ability to detect non-linguistic pitch differences, a measure that was obtained in a background experiment and entails the participants' accuracy at detecting height differences in pure tone pairs. Learners with good pitch perception abilities typically showed a negativity at left-lateral electrodes for the target minus control subtraction. As such, the five participants with best pitch perception accuracy (above 98%) had an amplitude of -0.19 μV at left electrode sites. Learners with inferior pitch perception abilities, on the other hand, showed no tendencies towards left-hemispheric negativities; notably, the five participants with lowest pitch perception accuracy (below 86%) had a positive-going response (with a mean amplitude of $+0.08$ μV) at left-lateral electrodes. This indicates that Non-Tonal L1s with particularly good pitch discrimination skills were more likely to exhibit an ELAN-like response for morphosyntactic activation, whereas on average their group did not. This result suggests that their relatively well-tuned auditory networks for pitch processing may allow even Non-Tonal L1s to automatically activate the novel words' morphosyntax in a manner similar to that of Tonal L1s. Thus, we believe that non-tonal L2 learners use their general pitch perception abilities to compensate for the lack of L1 word-level pitch processing systems that the tonal L1 learners can fall back on. If these pitch perception abilities are well developed, a non-tonal learner can tap into automatic processing at around 150 ms after the crucial input is perceivable and produce an ELAN, which we believe to signal the activation of morphosyntax.

The Non-Tonal L1s' LAN also differed from that elicited by the Tonal L1 group. While the Tonal L1s produced a LAN from very early on in the experiment, the Non-Tonal L1s needed the overnight consolidation period to do so. At the end of the first learning session, their response amplitude to both targets and controls was amplified as compared to the Tonal L1s'. 24 h later, i.e., after the consolidation stage, the Non-Tonal L1s' ERP amplitudes attenuated to the Tonal L1s' levels and a LAN for target words emerged. The most likely reason for the decreased ERP activity at the beginning of the second session for the Non-Tonal L1s is that the consolidated stimuli sounded less unexpected and therefore elicited a less pronounced neural response. Thus, for a group of participants with non-tonal background, morphosyntactic combinatorial processing of words with tonal morphemes is perhaps only possible after consolidation. This implies a slower process of memory formation and transfer, involving neocortical and hippocampal structures, unlike the immediate cortical signatures of learning manifest in Tonal L1s. We did, however, find a negative correlation of negativity at anterior electrodes on day 1 and response accuracy in the Non-Tonal L1 group. Learners at the higher end of the accuracy scale typically exhibited negativities at anterior electrodes for the target minus control contrast. The five top performing participants (accuracy above 91%), for instance, had an amplitude of -0.32 μV at anterior electrode sites. In comparison, learners at the lower end of the accuracy scale typically had positive amplitudes at anterior sites: the five participants at the bottom end of the accuracy scale (below 64%) had an amplitude of $+0.27$ μV at anterior electrodes. The correlation with accuracy suggests that non-tonal speakers who perform well in the experiment might not necessarily have to rely on consolidation but can integrate morphosyntax and semantics already on day 1. This results in the emergence of a LAN before overnight consolidation in good learners. As this late component is less automatic than the previous ones, it is conceivable that motivation to learn or focus can accelerate combinatorial rather than whole word processing for morphosyntax even in Non-Tonal L1s.

Finally, Non-Tonal L1s were indistinguishable from Tonal L1s in the P600 response. Both groups produced a P600 for the novel tonal words from the beginning to the end of the acquisition study. The P600 is renowned for being the morphosyntax-related component that is most frequently found in even low-proficient L2 learners and for learners that cannot rely on L1-L2 transfer (Steinhauer et al., 2009; van Hell and Tokowicz, 2010). The Non-Tonal L1s who have no L1 counterpart of the target language morphosyntactic tone can and do deploy relatively controlled whole word integration and revision to process novel tonal words, as reflected by the P600.

We suggest that the processing differences we found between Tonal L1s and Non-Tonal L1s are reflective of different types of processing (possibly due to differences in proficiency that remained undetected by the behavioural response) and emerge as a consequence of positive L1-L2 transfer of tone-morphosyntax pattern recognition processes in one group and a lack thereof in the other. In studies with second language learners, the P600 is usually the first component that is elicited for morphosyntactic processing (Steinhauer et al., 2009). As proficiency increases, the LAN and other early components emerge. In natural language acquisition, this most prominently happens for learners that benefit from positive transfer (van Hell and Tokowicz, 2010). Advocates of dual route processing models (e.g., Clahsen, 1999; Rodriguez-Fornells et al., 2001) consider the elicitation of the LAN to be related to combinatorial, rule-based processing of morphosyntax. Left-anterior negativities are found for instance for the formation of regular (presumably rule-based) but not irregular (presumably lexically stored) past tense in e.g., English (Newman et al., 2007), Swedish (Schremm et al., 2019) and German (Penke et al., 1997). Yet, both irregular as well as regular past tense formation normally elicit a P600, which is known for its role in more global, integrative processes and as such essential regardless of whether morphosyntactic features are stored as part of lexical items or decomposed as rules. Considering the second language learner proficiency findings in light of dual route models, it seems highly plausible that the two are related. Learners at mid proficiency levels, who are seen to elicit only P600s, are likely to store morphosyntactic information as part of a word's lexical entry. They would not separately activate morphosyntactic rules or properties or apply combinatorial rules, explaining the lack of ELAN and LAN responses. Highly proficient L2s, on the contrary, often aided by transfer, can decompose the morphosyntactic input which gives rise to rule-based processing. Given this argumentation, we believe that the Non-Tonal L1 learners in the present study, despite the simplicity of the stimuli, only reached a mid-proficiency level on day 1. They were able to process the morphosyntactic content of the novel words, visible in high accuracy levels and the P600 response, but initially stored each novel word as a separate lexical entry. On the second day of the study, after the novel words were consolidated, a LAN emerged in the Non-Tonal L1 group, suggestive of a higher proficiency level and the ability to combinatorially process the novel words on day 2. The Tonal L1 groups' processing, on the other hand, suggested high proficiency levels already on day 1, presumably owing to L1-L2 transfer. The Tonal L1s had morphosyntactically conditioned tones in their L1 which was believed to facilitate the acquisition of new words with morphosyntactic tone. Additionally, speakers with the same tonal L1 as our participants have been reported to use rule-based processing for their native tones in combination with pseudowords (Schremm et al., 2018). This is to say that the Tonal L1s not only had experience with morphosyntactically meaningful tones but also with combinatorial processes involving such tones. Said experience is likely reflected in specified neural networks, which we believe could have given our Tonal L1 learners a considerable advantage in the acquisition of the novel tonal words. The results are readily visible in the ERPs: The word recognition component at ~50 ms suggests that the Tonal L1s were able to quickly validate the word status of the tonal input, presumably because of the fact that their acoustic processing circuits were fine-tuned to the perception of tones. The early elicitation of the subsequent ERP effects, on the other hand, was most likely facilitated by the native tones' morphosyntactic connection. The neural networks responsible for the processing of morphosyntactic cues predicted by L1 tones and their rule-based application allowed the Tonal L1 participants to virtually instantaneously extract the rules associating the novel tones with morphosyntactic information. Once the morphosyntactic rules were deduced, an ELAN and a LAN response emerged, signalling the activation and combinatorial application of the morphosyntactic rule for the current input. In sum, it appears as though Non-Tonal L1s have a lower proficiency level and initially rely on whole-word access rather than rule application, while Tonal L1s range at high proficiency already at day 1 and are capable of rule-based combinatorial processing of the novel words' morphosyntactic properties. Correlations within the Non-Tonal L1 group suggest that subsets of the Non-Tonal L1 participants may arrive at combinatorial processing earlier than on day 2. Interestingly, although before the overnight consolidation period the tonal and non-tonal participants appeared to process the information differently at the neural level, the behavioural learning outcomes are identical. One potential reason for this could be ceiling effects due to the relative simplicity of the stimuli in the present study. This suggestion is corroborated by the fact that many previous behavioural studies on tone acquisition and the role of transfer found no differences between learners that did and did not have tone in their native language. Behavioural perception and learning outcome can be unaffected by transfer if the tonal contrasts are not overly complex and relatively easy to perceive, even though the underlying processing might differ. Whereas future studies are needed to verify this explanation, the present results clearly suggest that the same behavioural outcome is achieved through at least partially different neural mechanisms which depend on individual language background.

4. Conclusion

We used electroencephalographic recordings to examine whether rapid learning is possible even for complex linguistic material. Artificial tonal words with infixed morphosyntactic properties were taught to learner groups with tonal and non-tonal language backgrounds. Both groups showed rapid functional neural changes within the first 20 minutes of acquisition. To study possible benefits of positive L1 to L2 transfer, we analysed the groups separately and found that positive transfer allowed learners to instantaneously employ early as well as late neural processes, indicative of rule-based, combinatorial processing. Learners with a tonal L1 elicited an early word recognition response at 50 ms, session-initially, as well as an ELAN, LAN and P600 in response to the morphosyntactic content of the novel words. Non-Tonal L1s, in turn, could not profit from positive transfer. As a result, they elicited only a P600 on day 1, followed by a LAN after consolidation. We believe this to signal a shift between whole word processing on day 1 and combinatorial processing on day 2. Correlation analyses within the Non-Tonal group suggested that the elicitation of an ELAN for tonal morphosyntactic words is crucially dependent on auditory perception networks, and that the elicitation of the later LAN may rather be influenced by conscious effort. Despite the apparent difference in processing methods, behavioural performance was identical between learner groups. A study with more complex stimuli may lead to visible differences in behavioural factors as a result of the absence of early, automatic components in response to novel, tonal words in Non-Tonal L1s' speech processing.

5. Experimental procedure

5.1. Participants

Two participant groups, matched for gender, age, socioeconomic status and working memory span (assessed using an automated operation span task; Unsworth et al., 2005), participated in this study. The Tonal L1 group consisted of twenty-three native speakers of Swedish (13 females, mean age 23.7 years, $SD = 2.6$). The Non-Tonal L1 group consisted of twenty-three native speakers of German (12 females, mean age 23.7 years, $SD = 1.6$). All participants were right-handed as assessed by the revised Edinburgh Handedness Inventory (Williams, 2013) and had normal hearing defined as pure-tone hearing thresholds $\leq$ 20 dB Hearing Level (ISO, 2004). Background testing furthermore revealed that all participants were equally able to hear and correctly discern differences in non-linguistic pitch variations as well as durational variations in segments (i.e., vowel length).

Participants for the Non-Tonal L1 group were recruited from a pool of exchange students at Lund University. Despite having lived in Sweden for some months (M = 160 days, SD = 114), none of them had studied Swedish for more than 6 months (M = 70 days, SD = 52) and the reported weekly exposure to Swedish was low (M = 15 h, SD = 14.6, including lessons). The average self-assessed level of Swedish proficiency was A1 (=beginner) of the Common European Framework of Reference with the top of the group peaking at level B1 (=intermediate). Also, importantly, all but one of the participants reported to have never heard of Swedish word accents. Word accents were not part of any participant's vocabulary acquisition routines, nor was there any awareness of the tones' ties to grammatical suffixes.

A survey of the participants' language background revealed that none of the participants included in the analysis had previously learned any tone language before Swedish. Furthermore, the Non-Tonal L1 and Tonal L1 group were matched on the number of languages spoken at above elementary proficiency (M = 2, SD = 1). Yet, due to being exchange students abroad, the Non-Tonal L1 group had a temporary increase in their use of non-native languages ($M_{\text{Tonal L1s}}$ = 63.05%; $M_{\text{Non-Tonal L1s}}$ = 30.91%), predominantly English. The experiments were conducted in agreement with the ethical guidelines for experiments in the Declaration of Helsinki.

5.2. Stimuli

Word learning in the present study proceeded in the form of an auditory word – visual picture association paradigm with occasional questions containing feedback. Participants first heard an auditory presentation of a novel word or a pseudoword control followed by a meaning-assigning or non-meaning-assigning (scrambled) picture, cf. Fig. 10.

5.2.1. Sound stimuli

For the sound stimuli, we recorded four vowels (/a/ /ɛ/ /i/ /u/) as well as seven word initial consonants (/d/ /f/ /k/ /l/ /p/ /s/ /t/) and nine word final consonants (/f/ /k/ /l/ /m/ /n/ /p/ /r/ /s/ /t/) in an anechoic chamber. All sounds were uttered by a male native speaker of Russian, chosen as phonetically similar but distinct from both German and Swedish, to avoid differential carry-over effects (which would arise if the sounds were uttered by either German or Swedish speakers); notably the selected phonemes are present in all three languages. For the recordings, the initial and final consonants were followed or preceded, respectively, by a dummy vowel (/o/ or /ø/), which was later cut away using Praat (Boersma, 2001). All consonants and vowels were normalised to the same power, measured as root-mean square (RMS) amplitude. We then lengthened all initial consonants to 330 ms using initial silence where necessary. Vowels were standardised to 450 ms and all final consonants to 200 ms. We consequently spliced consonants and vowel with 10 ms transition phases into CVC syllables. In this way, 54 different pseudowords of one second length were constructed. For experimental reasons, word durations were somewhat longer than they would be in natural speech in both Swedish and German. In Swedish, long vowels in a focused CVC syllable would have an average duration of 246 ms, the final consonant averaging at 90 or 180 ms depending on voicing (Helgason et al., 2013). In German, long vowels in isolated words would have an average duration of 178 ms, the final consonant averaging at 79 or 130 ms depending on voicing (Braunschweiler, 1997). We deliberately constructed expressly long CVC words in order to ensure that the different consonants and vowels could be distinguished effortlessly. Even more importantly, an excessive vowel duration was chosen because a considerable amount of semantic and morphosyntactic information was disambiguated during the vowel and because it would make the pitch manipulations (see below) easy to perceive, even for speakers who were not sensitive to morphologically related pitch contours at the word level.

We tested the resulting pseudowords in a short perception study with eight native speakers of German and three native speakers of Swedish. In a self-paced listening paradigm, the participants were asked to write down the words, repeat them and report whether they reminded them of any existing words in their L1. Based on the results from the perception study, we selected 24 stimuli that were perceived uniformly as pseudowords and whose consonants and vowels were discerned most clearly, see Table 1.

By pitch-manipulation, four different tones (two level tones [high & low] and two contour tones [fall & rise]) were added onto each of the pseudowords. The high tone had a steady 138 Hz fundamental frequency (F0). The fall had an F0 of 138 Hz at the beginning of the vowel, and fell to 98 Hz during the vowel, where it remained for the second transition phase and the final consonant. The low tone had a steady pitch at 98 Hz. The rise started at 98 Hz at the beginning of the vowel, and rose to 138 Hz during the vowel, where it remained for the rest of the word, see Fig. 8. The F0 movements for the rise and fall were linear in order to get steady changes in pitch already at the very beginning of the vowel. This way, the disambiguation point for both vowel and tone, and thus the point at which the word could be discerned, coincided with the beginning of the vowel.

5.2.2. Picture stimuli

A total of 384 pictures were constructed to be used in the learning phase of the experiment. Half of the pictures showed people with 24 different professions. For a complete list of the professions, refer to Table 2. The other half of the pictures were based on the profession pictures where all pixels were randomly scrambled, to be used as a control. An example set of pictures for one profession can be seen in Fig. 9.

The pictures were constructed with the intention to clearly depict two different grammatical categories: gender, represented by either male (masculine) or female (feminine) workers, and number, illustrated by either one (singular) or multiple (plural) workers. The male and female workers on the pictures were identical in height and posture. The gender differences were expressed in hair style, facial hair, and/or clothing. For the number distinction, the singular pictures showed one worker while the plural pictures showed either two, three or four workers. The variety in plural pictures was chosen in order to avoid a "dual" or other exact numeric interpretation.

5.3. Experimental procedure

All subjects participated on two consecutive days. On both days, they were seated at a distance of one meter from a computer screen and asked to keep their left and right index finger on a response box on a table in front of them the entire time. The experiment was controlled by E-Prime 2 stimulation software (Psychology Software Tools Inc., Sharpsburg, PA). All auditory stimuli were routed through a GSI 16 Audiometer (Grason & Stadler Inc., Eden Prairie, MN) and presented at 70 dB SPL through a pair of circumaural earphones (California Headphone Company, Danville, CA). The presentation level was verified using a Brüel and Kjaer 2231 sound level meter with a 4134 microphone in a 4153 Artificial Ear.

Table 1
Pseudowords used in the main experiment.

Initial consonant	Vowels			
	a	ɛ	i	u
d	dap	dɛp	dif	duf
f	fap	fɛp	fif	fuf
k	kaf	kɛf	kip	kup
l	lap	lɛp	lir	lur
s	sap	sɛp	sis	sus
t	taf	tɛf	tip	tup

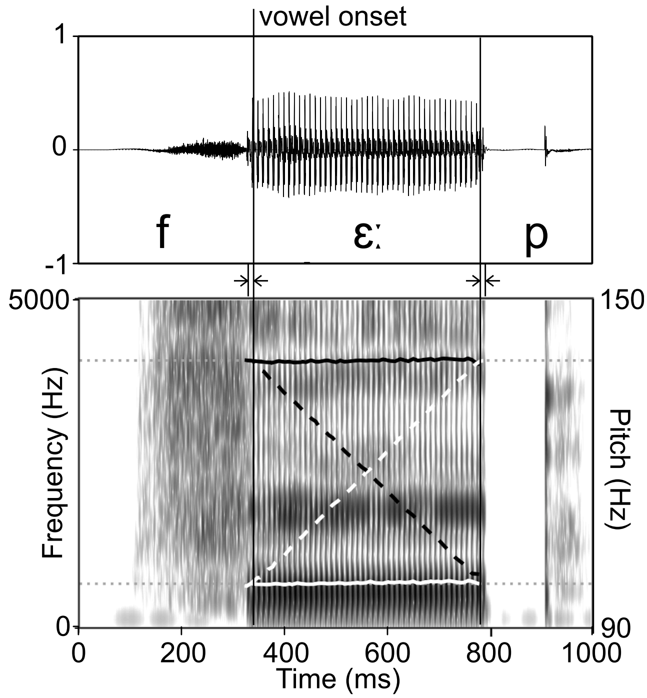

Fig. 8. Example of a stimulus word with acoustic waveform, spectrogram, and fundamental frequency in Hertz (Hz). The pitch (F0) contours show the four tonal variations: high (black continuous), fall (black dashed), low (white continuous), rise (white dashed). Two dotted light grey auxiliary lines are placed at 98 Hz and 138 Hz, respectively. Vowel onset and offset are marked by two vertical black lines. The position and length of the transition phases between segments are indicated between waveform and spectrogram.

Table 2
List of professions used for the picture stimuli.

Professions			
Ballet dancer	Doctor	Journalist	Flight attendant
Basketball player	Fire fighter	Painter	Tailor
Boxer	Fisher	Photographer	Teacher
Chemist	Flautist	Police officer	Tennis player
Cleaner	Gardener	Race driver	Violinist
Cook	Hairdresser	Singer	Waiter

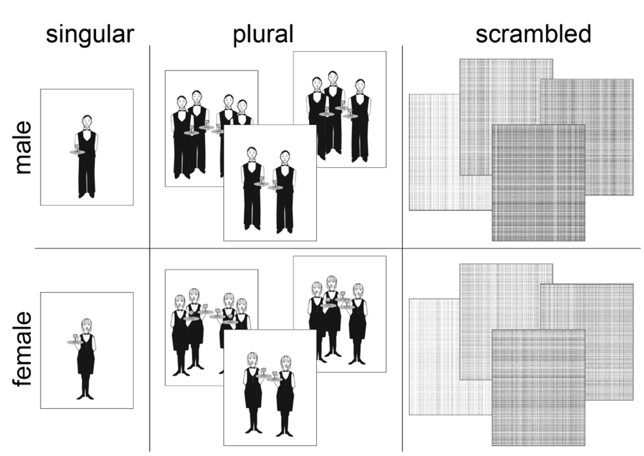

Fig. 9. Example of a set of picture stimuli used in the experiment.

5.3.1. *Instructions and training phase*

Day one of the experiment started with brief instructions on key behaviour during the ERP experiment, such as avoiding tension and keeping the index fingers on the response box throughout the recordings.

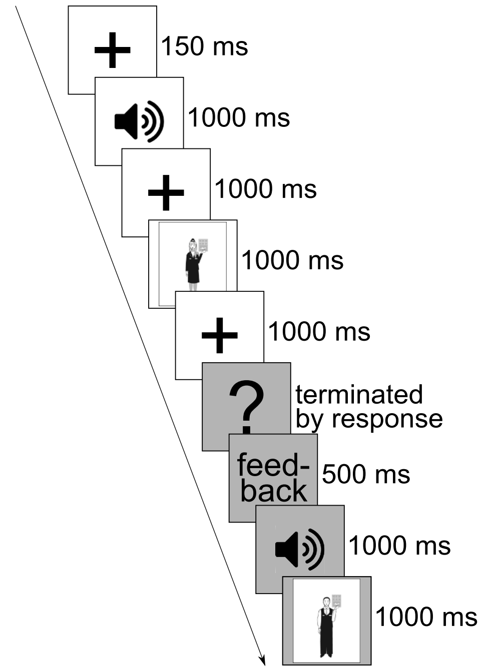

Fig. 10. Experimental procedure. Simple learning and control trials consisted of a sound stimulus followed by the corresponding picture stimulus (white slides). Question trials had an additional question slide, subsequent feedback and a repetition of the sound followed by the correct picture (grey slides).

Afterwards, participants were given written instructions. They were told that they were required to learn a number of words from a small, remote language. The words would denote professions and would be taught with the help of pictures. They were also told explicitly that number and gender in said language were expressed morphemically.

The written instructions on the first day were followed by a training phase, in which participants got acquainted with the learning procedure with the help of Spanish words (i.e., intentionally using a different language from the artificial one in the main experiment) with plural and gender distinctions. The training phase included two professions that were not used in the main experiment: architect(s) [*arquitecto/a(s)*] and mechanic(s) [*mecánico/a(s)*]. On day two, the detailed instructions and training phase were skipped. Instead, participants were briefly reminded of the main aspects of the study and key instructions for how to behave during the ERP experiment were repeated.

5.3.2. *Learning phase*

After the instructions and/or training, the main experiment started. Each participant was to learn 24 words, i.e., six different professions each with two genders and two numbers. The words' semantic content was pseudo-randomly chosen for each participant from the list of 24 professions and was always expressed through the initial and final consonant in the word (consonant frame). The two binary grammatical categories, gender and number, were expressed through differences in vowel (/a/ vs /ɛ/ or /i/ vs. /u/) or in tone (high vs. rise or low vs. fall). For an example, see Table 3. Every participant within one group had a unique combination of grammatical, segmental, and tonal features. The distribution of segmental and tonal features was counterbalanced, i.e., all four grammatical features were expressed equally often by all four vowels and four tones (i.e., three times each).[4]

On each day of the experiment, participants went through 30 learning cycles. In each cycle, there was a total of 51 trials: 24 learning trials, i.e., one trial for each to-be-learned word, as well as 24 control trials and three mismatch trials. Control trials differed from learning

[4] The exclusion of one participant (24 were initially recruited) in each nationality group resulted in a slight but insignificant dissymmetry.

Table 3
An example set of stimulus words for one participant. Professions are expressed within the words' consonant frame (here, *d_p* = chemist, *k_f* = tailor, etc.). Gender is signalled by either changes in vowel or tone (here: *ɛ* = feminine, *a* = masculine), and number by the other feature (here: high tone = singular, rising tone = plural). The control words have identical initial consonants and the complementing set of vowels and tones (here *i*, *u*, low tone, falling tone). They are presented with scrambled pictures and cannot be associated with meaning.

	Target words								Control words (no meaning)		
	Masculine				Feminine						
	Singular		*Plural*		*Singular*		*Plural*				
chemist	dap	*fall*	dap	*high*	dɛp	*fall*	dɛp	*high*	dif	/ duf	rise/low
firefighter	fap	*fall*	fap	*high*	fɛp	*fall*	fɛp	*high*	fif	/ fuf	rise/low
boxer	kaf	*fall*	kaf	*high*	kɛf	*fall*	kɛf	*high*	kip	/ kup	rise/low
ballet dancer	lap	*fall*	lap	*high*	lɛp	*fall*	lɛp	*high*	lir	/ lur	rise/low
waiter	sap	*fall*	sap	*high*	sɛp	*fall*	sɛp	*high*	sis	/ sus	rise/low
tailor	taf	*fall*	taf	*high*	tɛf	*fall*	tɛf	*high*	tip	/ tup	rise/low

trials in that the auditory stimuli contained a different set of words, i.e., the two tones that were not part of the target words, as well as the two non-target vowels and respective consonant frames. These were combined with scrambled pictures and served as controls for familiarisation as opposed to semantic and grammatical learning. In mismatch trials, a sound stimulus from the learned set was presented with a picture that was mismatched either in profession, gender or number.

In learning and control trials, a 1000-ms auditory stimulus was played followed by a 1000-ms fixation cross and a 1000-ms presentation of the corresponding stimulus picture. During the 1150-ms interstimulus interval (ISI), a fixation cross was present on the screen. The total stimulus onset asynchrony (SOA) for learning and control trials was thus 4150 ms, cf. slides with white background in Fig. 10.

The three mismatched trials as well as three pseudo-randomly selected learning trials were followed by a question slide prompting participants to decide whether sound and picture were matched correctly or not. Using their index fingers, participants indicated whether or not they had detected a mismatch. Half the participants used their right index finger for matched trials and their left index finger for mismatched trials; the other half had the opposite pattern. After the response, overt feedback was given and the trial's auditory stimulus was repeated, followed by the correct picture. After every 10 cycles (approximately 40 min), participants took a break to re-gain focus.

5.4. Electroencephalography (EEG)

The participants' brain activity was recorded using 64 Ag-AgCl EEG electrodes mounted in an electrode cap (EASYCAP GmbH, Herrsching, Germany), a SynAmps^2 EEG amplifier (Compumedics Neuroscan, Victoria, Australia), and Curry Neuroimaging Suite 7 software (Compumedics Neuroscan). To monitor eye movements, horizontal and vertical bipolar electrooculogram channels (EOG) were added. Impedances were kept below 3 kΩ for the scalp channels and below 10 kΩ for the eye channels. The left mastoid (M1) was used as online reference and the frontocentral electrode AFz as ground. EEG was recorded with a 500 Hz sampling rate using DC mode and an online anti-aliasing low pass filter at 200 Hz.

The recorded EEG data was then re-referenced offline to average reference, and subsequently filtered with a 0.01 Hz high pass and a 30 Hz low pass filter. ERP epochs of 1200 ms including a 200-ms baseline were extracted at the disambiguation point (i.e., vowel onset) for all auditory stimuli. Afterwards, an independent component analysis (ICA) (Jung et al., 2000) was conducted, and components representing eye artefacts and single bad channels were removed. Finally, all epochs still exceeding ± 100 μV were discarded.

5.5. Statistical analysis

A behavioural analysis was conducted separately for the two experimental factors 'Response Time' (RT) and 'Response Accuracy' which were recorded for question trials. The measures were submitted to a repeated measures Analysis of Variance (rmANOVA) with within-subject temporal factors 'Day' (day 1 vs. day 2) and 'Block' (early vs. late). The Tonal L1 and Non-Tonal L1 group were analysed separately in IBM SPSS Statistics 25 (International Business Machines Corp., Armonk, NY, United States) and Greenhouse-Geisser correction was used when applicable. Two-tailed independent samples t-tests were conducted to compare groups.

For the ERP analysis, time windows for effects were selected on the basis of global RMS peaks and in accordance with pre-existing literature. For the two early short gRMS peaks, 20-ms time windows were established around the peaks at 50–70 ms and 145–165 ms. For the relevant late and more wide-spread gRMS peaks, 100 ms time windows were analysed around the peaks: 460–560 ms and 670–770 ms.

In order to investigate effects of learning, mean ERP amplitudes for each group (i.e., Tonal L1s and Non-Tonal L1s) were submitted to an rmANOVA with the experimental factor 'Word Type' (target vs. control) and the temporal factors 'Day' (day 1 vs. day 2) and 'Block' (early vs. late) as well as the topographical distribution factors 'Laterality' (left, mid, right) and 'Posteriority' (anterior, central, posterior). The topographical factors correspond to the following electrode groups: left anterior with electrodes F7, F5, F3, FC5, and FC3; left central with C5, C3, CP5, CP3, and TP7; left posterior with P7, P5, P3, P07, and O1; mid anterior with F1, F2, FC1, FCZ, and FC2; mid central with C1, CZ, C2, CP1, and CP2; mid posterior with P1, PZ, P2, POz, and OZ; right anterior with F8, F6, F4, FC6, and FC4; right central with C6, C4, CP6, CP4, and TP8; and right posterior with P8, P6, P4, PO8, and O2. In time windows where it appeared as though overnight consolidation led to the emergence of significant effects, we made a separate analysis for the experimental factor 'Consolidation' (before vs. after) as well as the factors 'Word Type', 'Laterality' and 'Posteriority' (see above). The pre-consolidation condition consisted of the late trials on day 1, while the post-consolidation condition consisted of early trials on day 2. Two-tailed independent samples t-tests were used to compare groups.

We conducted one-tailed Pearson correlations to examine within-group variation which we believed to be related to non-significant results for between-group t-tests. For the 145–165 ms time window, mean amplitudes for targets minus controls at left lateral electrodes (left anterior, left central, left posterior) for the Non-Tonal L1 group were submitted to a correlation analysis with the behavioural variables 'Response Accuracy', 'Level of Swedish', 'Age of SLA Onset', and 'Accuracy in Non-Linguistic Pitch Distinction'. For non-linguistic pitch accuracy, participants completed a small background experiment where they judged the relative pitch of two piano tones. For the 460–560 ms time window on day 1, the same behavioural variables were submitted to a correlation analysis with mean amplitudes for targets minus controls at anterior electrodes (left anterior, mid anterior, right anterior) for the Non-Tonal L1s. As we did not have a priori expectations which of the measures would correlate with the difference in amplitude in the relevant topographical region, we applied Bonferroni corrections for the four comparisons run.

Acknowledgements

The work was supported by the Swedish Research Council (grant number 2011.2284), Knut and Alice Wallenberg Foundation (grant number 2014.0139), Danish Council for Independent Research (DFF 6110-00486, project 23776), RF Government (grant contract No. 14.W03.31.0010), and Marcus and Amalia Wallenberg Foundation (grant number 2014.0039). The authors would like to thank K. Jonas Brännström for his help with the auditory setup for the experiment, Frida Blomberg for her assistance during data acquisition and the Lund University Humanities Lab, in particular Annika Andersson, for providing a helpful and pleasant research environment.

Appendix A The phoneme inventory of German and Swedish. Slight qualitative vowel and consonant differences between the languages are highlighted by bold font. Vowels and consonants that are unique to one of the languages are underlined

Vowel inventory				Consonant inventory			
Long		Short		Voiceless		Voiced	
German	Swedish	German	Swedish	German	Swedish	German	Swedish
iː	iː	ɪ	ɪ	p	p	b	b
yː	yː	ʏ	ʏ	t	t	d	d
eː	eː	**ɛ**	**ɛ̣**	k	k	g	g
ɛː	ɛː			f	f	v	v
øː	øː	**œ**	**ø**	s	s	z	
aː	**ɑː**	a	a	**ʃ**	**ʂ**	**ʒ**	**ɕ**
oː	oː	ɔ	ɔ	ç		ʝ	ʝ
uː	uː	ʊ	ʊ	x		ɣ	
	ʉ̟ː		ɵ	χ		ʁ	
		ə		m	m		
		ɐ		n	n		
				ŋ	ŋ		
				l	l		
				ʀ	**r**		
						h	h
						ʔ	

References

Alves, M., 1995. Tonal features and the development of Vietnamese tones. *Working Papers*. Linguistics 27, 1–13.

Bakker, I., MacGregor, L.J., Pulvermüller, F., Shtryov, Y., 2013. Past tense in the brain's time: Neurophysiological evidence for dual-route processing of past-tense verbs. NeuroImage 71, 187–195. https://doi.org/10.1016/j.neuroimage.2012.12.065.

Banti, G., 1989. Two Cushitic systems: Somali and Oromo nouns. In: van der Hülst, H., Smith, N. (Eds.), Autosegmental studies on pitch accent. Foris Publications, Dordrecht, pp. 11–50.

Bloomfield, L., 1933. Language. Holt Rinehart Winston, New York.

Boersma, P., 2001. Praat, a system for doing phonetics by computer. Glot Int. 5 (9/10), 341–345.

Bou-Franch, P., 2013. Pragmatic transfer. In: Chapelle, C.A. (Ed.), The Encyclopedia of Applied Linguistics. Blackwell Publishing Ltd., pp. 1–5. https://doi.org/10.1002/9781405198431.wbeal0932.

Braun, B., Johnson, E.K., 2011. Question or tone 2? How language experience and linguistic function guide pitch processing. J. Phonet. 39, 585–594. https://doi.org/10.1016/j.wocn.2011.06.002.

Braunschweiler, N., 1997. Integrated cues of voicing and vowel length in German: A production study. Lang. Speech 40 (4), 353–376. https://doi.org/10.1177/002383099704000403.

Bruce, G., 1977. Swedish word accents in sentence perspective. Gleerup, Lund.

Charters, H., Dao, L., Jansen, L., 2012. Think of a number: conceptual transfer in second language acquisition of English plural-marking. CogniTextes 8. https://doi.org/10.4000/cognitextes.527.

Chen, A., Liu, L., Kager, R., 2015. Cross-linguistic perception of Mandarin tone sandhi. Lang. Sci. 48, 62–69. https://doi.org/10.1016/j.langsci.2014.12.002.

Clahsen, H., 1999. Lexical entries and rules of language: A multidisciplinary study of German inflection. Behav. Brain Sci. 22 (6), 991–1060. https://doi.org/10.1017/S0140525X99002228.

Crysmann, B., 2015. Representing morphological tone in a computational grammar of Hausa. J. Lang. Model. 3 (2), 463–512. https://doi.org/10.15398/jlm.v3i2.126.

Davis, M.H., Di Betta, A.M., Macdonald, M.J., Gaskell, M.G., 2009. Learning and consolidation of novel spoken words. J. Cognit. Neurosci. 21 (4), 803–820. https://doi.org/10.1162/jocn.2009.21059.

Francis, A.L., Ciocca, V., Ma, L., Fenn, K., 2008. Perceptual learning of Cantonese lexical tones by tone and non-tone language speakers. J. Phonet. 36, 268–294. https://doi.org/10.1016/j.wocn.2007.06.005.

Friederici, A.D., 2002. Towards a neural basis of auditory sentnece processing. TRENDS Cogn. Sci. 6 (2), 78–84. https://doi.org/10.1016/S1364-6613(00)01839-8.

Friederici, A.D., Pfeifer, E., Hahne, A., 1993. Event-related brain potentials duuring natural speech processing: Effects of semantic, morphological and syntactic violations. Cogn. Brain Res. 1, 183–192. https://doi.org/10.1016/0926-6410(93)90026-2.

Friederici, A.D., Steinhauer, K., Pfeifer, E., 2002. Brain signatures of artificial language processing: Evidence challenging the critical period hypothesis. PNAS 99 (1). https://doi.org/10.1073/pnas.012611199.

Gårding, E., 1998. Intonation in Swedish. In: Hirst, D., Di Cristo, A. (Eds.), Intonation systems: A survey of twenty languages. Cambridge University Press, Cambridge, pp. 112–130.

Ghazi-Said, L., Ansaldo, A.I., 2017. The neural correlates of semantic and phonological transfer effects: Language distance matters. Bilingualism: Language and Cognition 20 (5), 1080–1094. https://doi.org/10.1017/S136672891600064X.

Gibbon, D., 1998. Intonation in German. In: Hirst, D., Di Cristo, A. (Eds.), Intonation systems: A survey of twenty languages. Cambridge University Press, Cambridge, pp. 78–95.

Gillon Dowens, M., Vergara, M., Barber, H.A., Carreiras, M., 2009. Morphosyntactic processing in late second-language learners. J. Cognit. Neurosci. 22 (8), 1870–1887. https://doi.org/10.1162/jocn.2009.21304.

Gullberg, M., Indefrey, P. (Eds.), 2010. The earliest stages of language learning. Language Learning Research Club, University of Michigan.

Gullberg, M., Roberts, L., Dimroth, C., Veroude, K., Indefrey, P., 2010. Adult language learning after minimal exposure to an unknown natural language. Lang. Learn. 60 (S2), 5–24. https://doi.org/10.1111/j.1467-9922.2010.00598.x.

Gullberg, M., Roberts, L., Dimroth, C., 2012. What word-level knowledge can adult learners acquire after minimal exposure to a new language? Int. Rev. Appl. Linguist. Lang. Teach. 50, 239–276. https://doi.org/10.1515/iral-2012-0010.

Hagoort, P., 2003. How the brain solves the binding problem for language: A neurocomputational model of syntactic processing. NeuroImage 20, 18–29. https://doi.org/10.1016/j.neuroimage.2003.1796.

Hahne, A., 2001. What's different in second-language processing? Evidence from event-related brain potentials. J. Psycholinguist. Res. 30 (3), 251–266. https://doi.org/10.1023/A:1010490917575.

Hahne, A., Friederici, A.D., 1999. Electrophysiological evidence for two steps in syntactic analysis: Early automatic and late controlled processes. J. Cognit. Neurosci. 11 (2), 194–205. https://doi.org/10.1162/089892999563328.

Hahne, A., Friederici, A.D., 2001. Processing a second language: late learners' comprehension mechanisms as revealed by event-related brain potentials. Bilingualism: Language and Cognition 4 (2), 123–141. https://doi.org/10.1017/S1366728901000232.

Hahne, A., Friederici, A.D., 2002. Differential task effects on semantic and syntactic processes as revealed by ERPs. Cogn. Brain Res. 13 (3), 339–356. https://doi.org/10.1016/S0926-6410(01)00127-6.

Hallé, P.A., Chang, Y.-C., Best, C.T., 2004. Identification and discrimination of Mandarin Chinese tones by Mandarin Chinese vs. French listeneres. J. Phonet. 32, 395–442. https://doi.org/10.1016/S0095-4470(03)00016-0.

Hanna, J., Shtyrov, Y., Williams, J., Pulvermüller, F., 2016. Early neurophysiological indices of second language learning. Neuropsychologia 82, 18–30. https://doi.org/10.1016/j.neuropsychologia.2016.01.001.
Hao, Y.-C., 2012. Second language acquisition of Mandarin Chinese tones by tonal and non-tonal language speakers. J. Phonet. 40, 269–279. https://doi.org/10.1016/j.wocn.2011.11.001.
Hasting, A.S., Kotz, S.A., Friederici, A.D., 2007. Setting the stage for automatic syntax processing: The mismatch negativity as an indicator of syntactic priming. J. Cogn. Neurolingusit. 19 (3), 386–400. https://doi.org/10.1162/jocn.2007.19.3.386.
Hawkins, R., Lozano, C., 2006. Second language acquisition of phonology, morphology and syntax. In: Brown, K. (Ed.), The Encyclopedia of English LAnguage and Linguistics, 2nd ed. Elsevier, London, pp. 67–74. https://doi.org/10.1016/B0-08-044854-2/00634-9.
Hed, A., Schremm, A., Horne, M., Roll, M., 2019. Neural correlates of second language acquisition of tone-grammar associations. Mental Lexicon 14 (1), 97–121.
Helgason, P., Ringen, C., Suomi, K., 2013. Swedish quantity: Central Standard Swedish and Fenno-Swedish. J. Phonet. 41, 534–545. https://doi.org/10.1016/j.wocn.2013.09.005.
Isel, F., Hahne, A., Maess, B., Friederici, A.D., 2007. Neurodynamics of sentence interpretation: ERP evidence from French. Biol. Psychol. 74, 337–346. https://doi.org/10.1016/j.biopsycho.2006.09.003.
ISO, 2004. ISO 389–8. Acoustics: Reference zero for the calibration of audiometric equipment. Part 8: Reference equivalent threshold sound pressure levels for pure tone and circumaural earphones. International Organization for Standardization 389–8.
Jun, S.-A., Oh, M., 2000. Acquisition of second language intonation. The Journal of the Acoustical Society of America 106 (2802), 1–4. https://doi.org/10.1121/1.429024.
Jung, T.-P., Makeig, S., Humphries, C., Lee, T.-W., McKeown, M., Iragui, V., Sejnowski, T., 2000. Removing electroencephalographic artifacts by blind source separation. Psychophysiology 37, 136–178. https://doi.org/10.1111/1469-8986.3720163.
Kaan, E., 2014. Predictive sentence processing in L2 and L1: What is different? Linguistic Approaches to Bilingualism 4 (2), 257–282. https://doi.org/10.1075/lab.4.2.05kaa.
Kaan, E., Swaab, T.Y., 2003. Repair, revision, and complexity in syntactic analysis: An electrophysiological differentiation. J. Cognit. Neurosci. 15 (1), 98–110. https://doi.org/10.1162/089892903321107855.
Kasper, G., 1992. Pragmatic transfer. Second Language Research 8 (3), 203–231. https://doi.org/10.1177/026765839200800303.
Kimppa, L., Kujala, T., Leminen, A., Vainio, M., Shtyrov, Y., 2015. Rapid and automatic speech-specific learning mechanism in human neocortex. NeuroImage 118, 282–291. https://doi.org/10.1016/j.neuroimage.2015.05.098.
Kimppa, L., Kujala, T., Shtyrov, Y., 2016. Individual language experience modulates rapid formation of cortcial memory circuits for novel words. Sci. Rep. 1–10. https://doi.org/10.1038/srep30227.
Krott, A., Lebib, R., 2013. Electrophysiological evidence for a neural substrate of morphological rule application in correct wordforms. Brain Res. 1496, 70–83. https://doi.org/10.1016/j.brainres.2012.12.012.
Kuperberg, G.R., 2007. Neural mechanisms of language comprehension: Challenges to syntax. Brain Research(1146), 23–49. https://doi.org/10.1016/j.brainres.2006.12.063.
Kutas, M., Federmeier, K.D., 2014. Thirty years and counting: Finding meaning in the N400 component of the event related brain potential (ERP). Annu. Rev. Psychol. 62, 621–647. https://doi.org/10.1146/annurev.psych.093008.131123.
Kutas, M., Hillyard, S.A., 1980. Reading senseless sentences: Brain potentials reflect semantic incongruity. Science 207 (4427), 203–205. https://doi.org/10.1126/science.7350657.
Lado, R., 1957. Linguistics across cultures: Applied linguistics for language teachers. The University of Michigan, Ann Arbor Doi: 10.1017/S0008413100025196.
Le Gac, D., 2003. Tonal alternations in Somali. In: Lecarme, J. (Ed.), Research in Afroasiatic Grammar II. John Benjamins, Amsterdam, Philadelphia, pp. 287–304. https://doi.org/10.1075/cilt.241.15gac.
Lee, J.L., Matthews, S., 2015. When French becomes tonal: Prosodic transfer from L1 Cantonese and L2 English. In: Levis, J., Mohammed, R., Qian, M., Zhou, Z. (Eds.), The 6th Annual Proceedings of the Pronunciation in Second Language Learning and Teaching Conference, Santa Barbara, CA. Iowa State University, Ames, IA, pp. 63–72.
Lee, Y.-S., Vakoch, D.A., Wurm, L.H., 1996. Tone Perception in Cantonese and Mandarin: A cross-linguistic comparison. J. Psycholinguist. Res. 25 (5), 527–542. https://doi.org/10.1007/BF01758181.
Leminen, A., Kimppa, L., Leminen, M.M., Lehtonen, M., Mäkelä, J.P., Shtyrov, Y., 2016. Acquisition and consolidation of novel morphology in human neocortex: A neuromagnetic study. Cortex 83, 1–16. https://doi.org/10.1016/j.cortex.2016.06.020.
Luk, Z.P.-S., Shirai, Y., 2009. Is the acquisition order of grammatical morphemes impervious to L1 knowledge? Evidence from the acquisition of plural -s, articles, and possessive 's. Lang. Learn. 59 (4), 721–754. https://doi.org/10.1111/j.1467-9922.2009.00524.x.
MacGregor, L.J., Pulvermüller, F., van Casteren, M., Shtyrov, Y., 2012. Ultra-rapid access to words in the brain. *Nature*. Communications 3(711), 1, 7. https://doi.org/10.1038/ncomms1715.
Maddieson, I., 2013. Tone. In: Dryer, M.S., Haspelmath, M. (Eds.), The world atlas of language structures online. Max Planck Institute for Evolutionary Anthropology, Leipzig Retrieved September 17, 2018, from http://wals.info/chapter/13.
Mampe, B., Friederici, A.D., Christophe, A., Wermke, K., 2009. Newborns' cry melody is shaped by their native language. Curr. Biol. 19, 1994–1997. https://doi.org/10.1016/j.cub.2009.09.064.
McLaughlin, J., Osterhout, L., Kim, A., 2004. Neural correlates of second-language word learning: minimal instruction produces rapid change. Nat. Neurosci. 7 (7), 703–704. https://doi.org/10.1038/nn1264.
McPherson, L.E., 2014. *Replacive grammatical tone in the Dogon languages.* (doctoral dissertation). University of California, Los Angeles.
Meisel, J.M., 2009. Second language acquisition in early childhood. Zeitschrift für Sprachwissenschaft 28, 5–34. https://doi.org/10.1515/ZFSW.2009.002.
Molinaro, N., Barber, H.A., Caffarara, S., Carreiras, M., 2014. On the left anterior negativity (LAN): The case of morphosyntactic agreement. Cortex 66, 156–159. https://doi.org/10.1016/j.cortex.2014.06.009.
Neville, H., Nicol, J.L., Barss, A., Forster, K.I., Garret, M.F., 1991. Syntactically based sentence processing classes: Evidence from event-related potentials. J. Cognit. Neurosci. 3 (2), 151–165. https://doi.org/10.1162/jocn.1991.3.2.151.
Newman, A.J., Ullman, M.T., Pancheva, R., Waligura, D.L., Neville, H.J., 2007. An ERP study of regular and irregular English past tense inflection. NeuroImage 34, 435–445. https://doi.org/10.1016/j.neuroimage.2006.09.007.
Orie, O.O., 2006. L2 acquisition and Yoruba tones: Issues and challenges. In: Arasanyin, O.F., Pemberton, M.A. (Eds.), Selected Proceedings of the 36th Annual on African Linguistics. Cascadilla Proceedings Project, Sommerville, MA, pp. 121–128.
Osterhout, L., Holcomb, P.J., 1992. Event-related brain potentials elicited by syntactic anomaly. J. Mem. Lang. 31, 785–806. https://doi.org/10.1016/0749-596X(92)90039-Z.
Partanen, E., Leminen, A., de Paoli, S., Bundgaard, A., Skjold Kingo, O., Krøjgaard, P., Shtyrov, Y., 2017. Flexible, rapid and automatic neocortical word form acquisition mechanism in childen as revealed by neuromagnetic brain response dynamics. NeuroImage 450–459. https://doi.org/10.1016/j.neuroimage.2017.03.066.
Penke, M., Weyerts, H., Gross, M., Zander, E., Münte, T., Clahsen, H., 1997. How the brain processes complex words: An event-related potential stuy of German verb inflections. Cogn. Brain Res. 6 (1), 37–52. https://doi.org/10.1016/S0926-6410(97)00012-8.
Pulvermüller, F., Assadollahi, R., 2007. Grammar or serial order?: Discrete combinatorial brain mechanisms reflected by the syntactic mismatch negativity. J. Cognit. Neurosci. 19 (6), 971–980. https://doi.org/10.1162/jocn.2007.19.6.971.
Pulvermüller, F., Shtyrov, Y., 2003. Automatic processing of grammar in the human brain as revealed by the mismatch negativity. NeuroImage 20, 159–172. https://doi.org/10.1016/S1053-8119(03)00261-1.
Ramirez, G., Chen, X., Pasquarella, A., 2013. Cross-Linguistic Transfer of Morphological Awareness in Spanish-Speaking English Language Learners: The Facilitating Effect of Cognate Knowledge. Topics in Language Disorders 33 (1), 73–92. https://doi.org/10.1097/TLD.0b013e318280f55a.
Rast, R., 2010. The role of linguistic input in the first hours of adult language learning. Lang. Learn. 60 (Suppl. 2), 64–84.
Riad, T., 1998. The origin of Scandinavian tone accents. Diachronica 15 (1), 63–98. https://doi.org/10.1075/dia.15.1.04ria.
Riad, T., 2014. The phonology of Swedish. Oxford University Press, Oxford.
Rischel, J., 1963. Morphemic tone and word tone in Eastern Norwegian. Phonetica 10, 154–164. https://doi.org/10.1159/000258166.
Rodriguez-Fornells, A., Clahsen, H., Lleó, C., Zaake, W., Münte, T.F., 2001. Event-related brain responses to morphological violations in Catalan. Cogn. Brain Res. 11 (1), 47–58. https://doi.org/10.1016/S0926-6410(00)00063-X.
Roll, M., Söderström, P., Mannfolk, P., Shtyrov, Y., Johansson, M., van Westen, D., Horne, M., 2015. Word tones cueing morphosyntactic structure: Neuroanatomical substrates and activation time-course assessed by EEG and fMRI. Brain Lang. 150, 14–21. https://doi.org/10.1016/j.bandl.2015.07.009.
Roll, M., Söderström, P., Frid, J., Mannfolk, P., Horne, M., 2017. Forehearing words: Pre-activation of word endings at word onset. Neurosci. Lett. 658, 57–61. https://doi.org/10.1016/j.neulet.2017.08.030.
Rothman, J., Cabrelli Amaro, J., 2010. What variables condition syntactic transfer? A look at the L3 initial state. Second Language Research 26 (2), 189–218. https://doi.org/10.1177/0267658309349410.
Sainz, J.S., Lazaro, M., 2009. Early and late effects of morphological decomposition: Brain correlates of family size effects on complex words and pseudowords. Proceedings of the Annual Meeting of the Cognitive Science Society 31, 2992–2997.
Schremm, A., Hed, A., Horne, M., Roll, M., 2017. Training predictive L2 processing with a digital game: Protoype promotes acquisition of anticipatory use of tone-suffix associations. Comput. Educ. 114, 206–221. https://doi.org/10.1016/j.compedu.2017.07.006.
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 Lang. 176, 42–47. https://doi.org/10.1016/j.bandl.2017.12.001.
Schremm, A., Novén, M., Horne, M., Roll, M., 2019. Brain responses to morphologically complex verbs: An electrophysiological study of Swedish regular and irregular past tense forms. Journal of Neurolinguistics 51, 76–83. https://doi.org/10.1016/j.jneuroling.2019.01.006.
Shen, G., Froud, K., 2016. Categorical perception of lexical tones by English learners of Mandarin Chinese. The Journal of the Acoustical Society of America 140 (6), 4396–4403. https://doi.org/10.1121/1.4971765.
Shih, C., Lu, H.-Y.D., 2010. Prosody transfer and suppression: Stages of tone acquisition. Speech Prosody 2010, 968.
Shtyrov, Y., Lenzen, M., 2017. First-pass neocortical processing of spoken language takes only 30 msec: Electropysiological evidence. Cognitive Neuroscience 8 (1), 24–38. https://doi.org/10.1080/17588928.2016.1156663.
Shtyrov, Y., Pulvermüller, F., Näätänen, R., Ilmoniemi, R.J., 2007. Grammar processing outside the focus of attention: an MEG study. J. Cognit. Neurosci. 15 (8), 1195–1206. https://doi.org/10.1162/089892903322598148.
Shtyrov, Y., Nikulin, V.V., Pulvermüller, F., 2010. Rapid cortical plasticity underlying novel word learning. J. Neurosci. 30 (50), 16864–16867. https://doi.org/10.1523/JNEUROSCI.1376-10.2010.
Simon, E., 2010. Phonological transfer of voicing and devoicing rules: Evidence from L1 Dutch and L2 English conversational speech. Lang. Sci. 32, 63–86. https://doi.org/10.1016/j.langsci.2008.10.001.

Sneed German, E., Herschensohn, J., Frenck-Mestre, C., 2015. Pronoun processing in anglophone late L2 learners of French: Behavioral and ERP evidence. Journal of Neurolinguistics 40, 15–40. https://doi.org/10.1016/j.jneuroling.2014.12.001.

So, C.K., Best, C.T., 2010. Cross-language perception of non-native tonal contrasts: Effects of native phonological and phonetic influences. Lang. Speech 53 (2), 273–293. https://doi.org/10.1177/0023830909357156.

Söderström, P., Horne, M., Frid, J., Roll, M., 2016. Pre-activation negativity (PrAN) in brain potentials to unfolding words. Front. Hum. Neurosci. 10 (512), 1–11. https://doi.org/10.3389/fnhum.2016.00512.

Steinhauer, K., 2014. Event-related potentials (ERPs) in second language research: A brief introduction to the technique, a selected review, and an invitation to reconsider critical periods in L2. Applied Linguistics 35 (4), 393–417. https://doi.org/10.1093/applin/amu028.

Steinhauer, K., Drury, J.E., 2012. On the early left-anterior negativity (ELAN) in syntax studies. Brain Lang. 120, 135–162. https://doi.org/10.1016/j.bandl.2011.07.001.

Steinhauer, K., White, E.J., Drury, J.E., 2009. Temporal dynamics of late second language acquisition: evidence from event-related brain potentials. Second Langauge Research 25 (1), 13–41. https://doi.org/10.1177/0267658308098995.

Tokowicz, N., MacWhinney, B., 2005. Implicit and explicit measures of sensitivity to violations in second language grammar: An event-related potential investigation. Studies in Second Language Acquisition 27, 173–204. https://doi.org/10.1017/S0272263105050102.

Tronnier, M., Zetterholm, E., 2015. Appropriate tone accent production in L2-Swedish by L1-speakers of Somali? In: Proceedings of the International Symposium on the Acquisition of Second Language Speech. 5, pp. 722–736 Concordia Working Papers in Applied Linguistics.

Tsimpli, I.M., 2014. Early, late or very late? Timing acquisition and bilingualism. Linguistic Approaches to Bilingualism 4 (3), 283–313. https://doi.org/10.1075/lab.4.3.01tsi.

Unsworth, N., Heitz, R.P., Schrock, J.C., Engle, R.W., 2005. An automated version of the operation span task. Behavior Research Methods 37 (3), 498–505. https://doi.org/10.3758/BF03192720.

van Dommelen, W.A., Husby, O., 2007. The perception of Norwegian word tones by Chinese and German listeners. In: New Sounds 2007: Proceedings of the Fifth International Symposium on the Acquisition of Second Language Speech, pp. 165–174.

van Hell, J.G., Tokowicz, N., 2010. Event-related brain potentials and second language learning: syntactic processing in late L2 learners at different L2 proficiency levels. Second Language Research 26 (1), 43–74. https://doi.org/10.1177/0267658309337637.

Wang, Y., Spence, M.M., Jongman, A., Sereno, J.A., 1999. Training American listeneres to perceive Mandarin tones. The Journal of the Acoustical Society of America 106 (3649), 1–10. https://doi.org/10.1121/1.428217.

Wang, Y., Jongman, A., Sereno, J.A., 2001. Dichotic perception of Mandarin tones by Chinese and American listeners. Brain Language 78, 332–348. https://doi.org/10.1006/brln.2001.2474.

Wang, Y., Jongman, A., Sereno, J.A., 2003. Acoustic and perceptual evaluation of Mandarin tone productions before and after perceptual training. J. Acoust. Soc. Am. 113 (2), 1–11. https://doi.org/10.1121/1.1531176.

Wayland, R.P., Li, B., 2008. Effects of two training procedures in cross-language perception of tones. J. Phonet. 36, 250–267. https://doi.org/10.1016/j.wocn.2007.06.004.

Weber-Fox, C.M., Neville, H.J., 1996. Maturational constraints on functional specializations for langauge processing: ERP and behavioral evidence in bilingual speakers. J. Cognit. Neurosci. 8 (3), 231–256. https://doi.org/10.1162/jocn.1996.8.3.231.

Wiese, R., 1996. The phonology of German. Oxford University Press, Oxford.

Williams, S.M., 2013. A major revision of the Edinburgh Handedness Inventory. https://doi.org/10.13140/2.1.5180.9284.

Yang, C., Chan, M.K., 2010. The perception of mandarin chinese tones and intonation by American learners. Journal of the Chinese Language Teachers Association 45 (1), 7–36.

Yip, M.J., 1980. *The tonal phonology of Chinese.* (doctoral dissertation). Institute of Technology, Massachusetts.

Yip, M.J., 2002. Tone. Cambridge University Press, Cambridge doi: 10.1017/CBO9781139164559.

Zetterholm, E., Tronnier, M., 2012. The pronunciation of Somali-accented Swedish. In: FONETIK 2012. Department of Philosophy, Linguistics and Theory of Science, University of Gothenburg, pp. 93–96.