Rapid neural processing of grammatical tone in second language learners

Gosselke Berthelsen, Sabine

2021

Document Version:
Publisher's PDF, also known as Version of record

Link to publication

Citation for published version (APA):
Gosselke Berthelsen, S. (2021). *Rapid neural processing of grammatical tone in second language learners*. [Doctoral Thesis (compilation), Centre for Languages and Literature]. Lund University.

Total number of authors:
1

Rapid neural processing of grammatical tone in second language learners

SABINE GOSSELKE BERTHELSEN
CENTRE FOR LANGUAGES AND LITERATURE | LUND UNIVERSITY

Rapid neural processing of grammatical tone in second language learners

Learning a second language can feel like a slow and tedious process. Yet, the human brain starts segmenting and categorising new language input virtually from the very second that it is exposed it. It supposedly does so based on what it currently knows about how languages sound and how they are structured. The present dissertation investigates the suggested importance of familiarity with the properties of a new language in early stages of language learning. It focuses primarily on the feature of grammatical tone, that is, a systematic change in pitch associated with grammatical function. For example, different pitch patterns might be related to singular and plural interpretations of nouns. In four papers, the thesis examines the acquisition of grammatical tone both in a natural language (Swedish) and in strongly controlled, artificial words. With the help of electroencephalography (EEG), it explores how learners from a tonal and a non-tonal language background process the novel tonal information in correct and mismatched contexts during learning. Analysing different processing steps related to varying degrees of automaticity, the present dissertation shows how the processing of foreign words is affected by the similarity between native and foreign language, how tone influences the speed of second language processing, and what role mismatches (or violations) play in this context. The findings underline the importance of familiarity with the sounds of a new language, strengthen the proposition that second language tone might generally be difficult to process, and question the use of violation paradigms when studying second language learners. The dissertation contributes to the understanding of how quickly second language learners process new language input and how the automatic segmentation of novel input crucially depends on familiarity with the new language's properties.

Printed by Media-Tryck, Lund 2021 NORDIC SWAN ECOLABEL 3041 0903

LUND UNIVERSITY

Lund University
The Faculties of Humanities and Theology
Centre for Languages and Literature
ISBN 978-91-89213-38-8

Rapid neural processing of grammatical tone in second language learners

Sabine Gosselke Berthelsen

LUND UNIVERSITY

DOCTORAL DISSERTATION

by due permission of the Faculties of Humanities and Theology,
Lund University, Sweden.
To be defended at LUX C121, Friday 5 February 2021, 10.15-12.00.

Faculty opponent
Antoni Rodríguez-Fornells
Catalan Institution for Research and Advanced Studies,
ICREA, Barcelona

Organization
LUND UNIVERSITY
Centre for Languages and Literature

Author Sabine Gosselke Berthelsen

Document name DOCTORAL DISSERTATION

Date of issue 5 February 2021

Sponsoring organization

Title and subtitle

Abstract
The present dissertation investigates how beginner learners process grammatical tone in a second language and whether their processing is influenced by phonological transfer. Paper I focuses on the acquisition of Swedish grammatical tone by beginner learners from a non-tonal language, German. Results show that non-tonal beginner learners do not process the grammatical regularities of the tones but rather treat them akin to piano tones. A rightwards-going spread of activity in response to pitch difference in Swedish tones possibly indicates a process of tone sensitisation. Papers II to IV investigate how artificial grammatical tone, taught in a word-picture association paradigm, is acquired by German and Swedish learners. The results of paper II show that interspersed mismatches between grammatical tone and picture referents evoke an N400 only for the Swedish learners. Both learner groups produce N400 responses to picture mismatches related to grammatically meaningful vowel changes. While mismatch detection quickly reaches high accuracy rates, tone mismatches are least accurately and most slowly detected in both learner groups. For processing of the grammatical L2 words outside of mismatch contexts, the results of paper III reveal early, preconscious and late, conscious processing in the Swedish learner group within 20 minutes of acquisition (word recognition component, ELAN, LAN, P600). German learners only produce late responses: a P600 within 20 minutes and a LAN after sleep consolidation. The surprisingly rapid emergence of early grammatical ERP components (ELAN, LAN) is attributed to less resource-heavy processing outside of violation contexts. Results of paper IV, finally, indicate that memory trace formation, as visible in the word recognition component at ~50 ms, is only possible at the highest level of formal and functional similarity, that is, for words with falling tone in Swedish participants. Together, the findings emphasise the importance of phonological transfer in the initial stages of second language acquisition and suggest that the earlier the processing, the more important the impact of phonological transfer.

Key words
grammatical tone, phonology, phonetics, word accents, morphology, inflection, grammar, EEG, ERP, processing, violations, ELAN, LAN, P600, SLA, second language acquisition, beginner learner, transfer, German, Swedish

Classification system and/or index terms (if any)

Supplementary bibliographical information

Language

ISSN

ISBN 978-91-89213-38-8

Recipient's notes

Number of pages 83

Price

Security classification

Signature

Date 2020-12-16

Rapid neural processing of grammatical tone in second language learners

Sabine Gosselke Berthelsen

LUND UNIVERSITY

Cover by Sabine Gosselke Berthelsen.

Faculties of Humanities and Theology
Centre for Languages and Literature

ISBN 978-91-89213-38-8 (print)
ISBN 978-91-89213-39-5 (digital)

Printed in Sweden by Media-Tryck, Lund University
Lund 2021

NORDIC SWAN ECOLABEL
Printed matter 3041 0903

Media-Tryck is an environmentally certified and ISO 14001:2015 certified provider of printed material. Read more about our environmental work at www.mediatryck.lu.se

MADE IN SWEDEN

To my parents and my grandmother
for teaching me to be curious

Table of Contents

Acknowledgements

There are many people whose continuous support and guidance have been vital throughout the journey that has led to this dissertation.

First and foremost, my deepest gratitude goes to my main supervisor Mikael Roll, who insistently nudged me to embark on this journey and who has since been a constant source of knowledge, inspiration, support, and encouragement. Thank you for not taking no for an answer, for teaching me the ropes, and for always believing in me. I also owe a great debt of gratitude to my co-supervisors Merle Horne and Yury Shtyrov. Merle's intellectual curiosity and enthusiasm have been truly infectious and inspiring and Yury's scientific versatility and overview have given my work much-needed focus and perspective. Thank you all for your patience and dedication over the years.

A sincere thank you goes to Line Burholt Kristensen for an inspiring discussion and insightful comments during and after the mock defense. Susan Sayehli, Nicolai Pharao, and Victoria Johansson, thank you for accepting to be members on my thesis defense committee. Finally, I extend a heartfelt thank you to Antoni Rodríguez-Fornells for agreeing to act as Faculty opponent.

For providing a tremendously stimulating, supportive and respectful research environment, a big thank you goes to all previous and current members of the Neurolinguistics group: Pelle Söderström, Frida Blomberg, Otto Ewald, Anna Hed, Andrea Schremm, Mikael Novén, Renata Kochancikaite, Andrea Fingerhut, and Anna Hjortdal. I am so glad to have been part of a group with such talent and integrity. Thank you also to all the current and former PhD students at the department, in particular Sandra Debreslioska, Malin Svensson Lundmark, Vi Than Son, and Briana van Epps. It has been a delight to work alongside you all.

I could not have carried out my research without Jonas Brännström's kind help with all things auditory. I also thank Annika Andersson for her help and guidance with respect to EEG and the Neuroscan system, and Marianne Gullberg for her interest in my work, her input, and advice. Thank you to Frida Blomberg for help during capping, for test-piloting the experiment, and for providing moral support, to Sandra Debreslioska for test-piloting and confidence building, and to both of them for invaluable feedback on a draft of this thesis.

I would like to thank all of my wonderful participants, especially those in the main study who were stuck with me in a tiny basement room for 12 hours. I also thank family and friends who helped with an evaluation of an early version of the auditory stimuli: Judith Mirschel, Katrin Heinermann, Marcel Deppenmeier, and Eva Berg.

Thank you to the Humanities Lab for providing an inspiring and enjoyable research environment with great facilities and kind staff and users.

I further owe great thanks to my Bachelor thesis supervisors who took me aside and gently pushed me ‘Richtung’ academia. Thank you Claudia Wich-Reif and Anette Rosenbach.

Finally, a sincere thank you to family and friends who have ground me throughout the years. A special thanks to my husband Brian for amazing support and to my wonderful little girls, Lærke and Freja, for putting up with a very sleep deprived and rather grumpy version of myself towards the end of the PhD.

List of papers

Paper I

Gosselke Berthelsen, S., Horne, M., Brännström, K. J., Shtyrov, Y., & Roll, M. (2018). Neural processing of morphosyntactic tonal cues in second-language learners. *Journal of Neurolinguistics*, *45*, 60-78. doi:10.1016/j.jneuroling.2017.09.001

I adapted the study and experimental procedure from Roll et al. (2015) to second language learners, acquired and analysed all data, and was the main author of the manuscript.

Paper II

Gosselke Berthelsen, S., Horne, M., Shtyrov, Y., & Roll, M. (2020b*)*. Phonological transfer effects in novice learners: A learner's brain detects grammar errors only if the language sounds familiar. Manuscript submitted.

I conceptualised the study, acquired and analysed all data, and was the main author of the manuscript.

Paper III

Gosselke Berthelsen, S., Horne, M., Shtyrov, Y., & Roll, M. (2020a). Different neural mechanisms for rapid acquisition of words with grammatical tone in learners from tonal and non-tonal backgrounds: ERP evidence. *Brain Research*, *1729*(146614). doi:10.1016/j.brainres.2019.146614

I conceptualised the study, acquired and analysed all data, and was the main author of the manuscript.

Paper IV

Gosselke Berthelsen, S.; Horne, M.; Shtyrov, Y.; & Roll, M. (2020c). Native language experience narrowly shapes pre-attentive foreign tone processing and governs rapid memory trace build-up: An ERP study. Manuscript submitted.

I conceptualised the study, acquired and analysed all data, and was the main author of the manuscript.

Abbreviations

Interlinear Glossing

2	second person
3	third person
DEF	definite
FEM	feminine
GEN	genitive
IMP	imperative
IND	indefinite
LINK	linking particle
MAS	masculine
NOM	nominative
PL	plural
PST	past
PTCP	participle
SG	singular
.	non-segmentable morpheme
-	segmentable morpheme

Syntax

NP	noun phrase
PP	prepositional phrase

Phonology

CVC	syllable consisting of consonant, vowel, and consonant
H	high tone
L	low tone
*	stressed syllable
$_1$	accent 1
$_2$	accent 2

Second language acquisition

SLA	second language acquisition
L1	native language
L2	second language
TL1	learners with tonal native language
NTL1	learners with non-tonal native language

Statistics

ANOVA	Analysis of Variance
d’	d-prime
F	F-ratio
FDR	false discovery rate
M	sample mean
Mdn	median
N	population size
n	sample size
p	p-value
r	Pearson’s correlation coefficient
RT	response times
SD	standard deviation
t	t-value

Neurophysiology

EEG	electroencephalography
ELAN	early left anterior negativity
ERP	event-related potential
gRMS	global root-mean-square
LAN	left anterior negativity
MEG	magnetoencephalography
MMN	mismatch negativity
ms	milliseconds
PrAN	pre-activation negativity
sMMN	syntactic mismatch negativity
µV	microvolt

1 Introduction

Humans have the fascinating ability of assessing novel situations with intriguing speed. New languages are no exception. Within weeks, days, or even just minutes, learners extract meaning and rules from new language input. Consider the sentence *Tim got two small dolls, six big, new books, and one old, red book for Christmas.* A learner of English, who has never heard the words *doll* and *book*, for instance, can deduce that they are things that can be given as Christmas presents and that can differ in size, age, or colour. This type of extracted information is related to the meaning of the words in the sentence and referred to as semantic. The same learner might also notice, that there is a systematic difference between *two dolls*, *six books* and *one book* such that an -*s* is added at the end of the word when the word is in plural. They might also notice, that descriptive adjectives like *blue* and *new* can appear before the nouns (*book*, *doll*). This type of extracted information is related to the structure of words and sentences and referred to as grammatical. However, there are also features in the sentence that an inexperienced learner might fail to notice, such as a fixed order of adjectives: size goes before age goes before colour. Thus, while learners can quickly deduce various semantic and grammatical features of novel language input, not all features may be equally accessible.

One type of feature that, supposedly, is particularly difficult to detect and categorise in a new language is the systematic use of pitch differences on individual words. Many languages around the world use pitch height or pitch movement to differentiate words from each other (lexical tone) or to add grammatical meaning to them (grammatical tone). Learners often struggle with categorising the information in the pitch dimension into distinctive, meaningful units. Mandarin Chinese, for instance, has one tone with rising pitch movement and one tone that starts with a small fall but ends in a rising pitch. The tones differentiate words such that, on the syllable *ba*, the rising tone can mean 'to pull up' while the fall-rise, amongst others, means 'to hold'. However, since both tones include a rising pitch movement, learners often experience difficulty in the distinction of the two. Swedish, on the other hand, has a grammatical tone system in which grammatical affixes are closely related to tones. Hence, the definite singular suffix -*en* is accompanied with a low tone on a noun stem, for example *bil* ('car'), while the plural suffix -*ar* requires the noun stem to be realised with a high tone.

Importantly, while tone learning is generally claimed to be challenging, some learners fare better in the distinction and, consequentially, the acquisition of tone than others do. The arguably most facilitative factor in this context is familiarity with the concept of tone from the native language. Thus, learners who have experience with listening out for different pitch cues, such as height or movement, in their native language are more adept at categorising these into potentially meaningful tone types in their second language. Their relatively more proficient tone distinction and acquisition abilities are likely to be based on strong and fine-tuned brain networks for the processing of pitch information. The networks can presumably also be activated for foreign tone which gives rise to tone detection and classification advantages for learners with tone in their native language in comparison to tone-naïve learners. For example, Swedish speakers would have an advantage when learning a new tone language, in comparison to German speakers, since German does not have word-level tone.

One way of detecting and measuring the acquisition of novel features in a second language is by means of recording the electrical activity in the brain with electroencephalography (EEG). The brain produces specific response patterns which are associated with the processing of different parts of language, such as meaning or grammatical rules. During the acquisition of a new language, learners absorb words and structures into their internal language processing system. This process is reflected in the emergence of language-related brain responses. With respect to the internalisation of new grammatical information, learners typically go through several consecutive acquisition stages, visible in the emergence of different brain responses. This makes it possible to monitor learners in their L2 acquisition process, observing how their brain responses become progressively more similar to those of native speakers.

The studies in the current thesis examine beginner learners' acquisition of grammatical tone in a second language context by recording electrical responses from the learner's brain using EEG. The studies focus on how long it takes for learners to produce typical neural responses for the processing of grammatical tone and whether previous familiarity with tone (i.e., because the learners speak a tonal native language) has a facilitative effect.

2 Background

2.1 Grammar and grammatical tone

Grammar, also referred to as morphosyntax, is an integral part of human language. It is a system of rules and regularities by which semantic and pragmatic meaning can be expressed. Such rules and combinatorial processes are found at the level of morphology, that is, regulating how morphemes are combined into words (Marantz, 1997; Marcus et al., 1995), and syntax, that is, structuring how words are combined into sentences (Chomsky, 1965). Morphemes are the smallest meaningful building blocks of language and can either be lexical (e.g., nouns or verbs) or grammatical (e.g., articles or affixes). Lexical morphemes can typically stand on their own (they are unbound) and express semantic concepts or meaning (*cat* = FELINE). Grammatical morphemes can either be unbound (e.g., articles in English) or bound (affixes) and have a grammatical function (*the* = DEFINITE, -*s* = PLURAL). Importantly, following the specific morphological rules in a given language, lexical and grammatical morphemes can be combined into words: *cats*. Words can further be combined into phrases (*the cats*) or sentences (*Curt got the cats for Christmas*.) according to syntactic word order rules. Like morphological rules, syntactic rules are language-specific and dictate, for instance, that articles, in English, be placed before the noun.

2.1.1 Nominal inflection

In the present dissertation, the primary focus lies on morphology, in particular on inflectional affixes on nouns. Inflectional affixes are grammatical morphemes that attach to nouns or verbs and modify them grammatically. The affixes that were used in the studies within this thesis relate to definiteness, number, and gender. In Swedish, the language studied in paper I, noun stems, such as *bil* ('car'), can be inflected by suffixes (affixes that appear after the word stem) for definiteness (*bil-en*, car-DEF.SG, 'the car') or number (*bil-ar*, car-IND.PL, 'cars'). The definite suffix is further specified for grammatical gender. In present-day Swedish, nouns have either neuter (neutrum: definite singular -*en*) or common (utrum: definite singular -*et*) gender. Although Swedish has lost the traditional distinction between feminine and masculine (Enger, 2005), a few words still have explicitly feminine forms with relatively transparent gender suffixes (e.g., *vän-inna*, friend-FEM, 'female friend' or *lärar-inna*, teacher-FEM, 'female teacher').

Besides suffixes, inflection can also be applied in the form of stem modulations (Bauer, 2003), such as changes in the segmental (vowels and consonants) or suprasegmental (e.g., stress or tone) features of a word. Thus, a change in the grammatical feature number can induce a change in the vowel of the English word *man*, that is, from the singular *man* to the plural *men*. Morphologically conditioned vowel changes on verbs and nouns are very common in the Germanic languages. This is due, in particular, to two systematic sound changes applied at different times in the history of the Germanic languages: the Indo-European ablaut and the more recent Germanic umlaut. While initially conditioned by morphology or phonology (combinatorial rules of sounds, see Wiese, 1997), umlaut and ablaut are largely fossilised in present-day languages (Féry, 1994; Riad, 2014). In addition to the segmental level, stem modifications can also affect the suprasegmental level where a grammatical inflection is realised within a change in the prosodic structure of a word. Thus, the stress on the English word *increase* compared to its homograph *increase* defines whether the word is a verb or a noun. Importantly, suprasegmental stem changes are also applied in the context of grammatical tone (see below), where grammatical function is expressed via changes in a word's pitch pattern. In the East Cushitic language Arbore, for instance, a change in the pitch of the word stem from non-high (*naag*, 'girl') to high (*náág*, 'boy') induces a change in grammatical gender (Banti, 1998).

2.1.2 Grammatical tone

Tone is a common feature in the world's languages. It is estimated that up to 70% of the world's languages have tone (Yip, 2002). Tone can be defined as the lexically or grammatically meaningful use of pitch. According to the function of tone, we discriminate between lexical tone and grammatical tone. Tone languages operate on a continuum where strong lexical and strong grammatical tone languages form the end points (Hyman, 2016). Lexical tone is prominent throughout the East Asian languages families, for instance, and distinguishes words: The Mandarin syllable *ma* produced with a high level tone (*má* [妈]) means 'mother' while the same syllable produced with a falling tone (*mâ* [蚂]) means 'grasshopper'. Grammatical tone is most prominent in different African language families and is related to grammatical function. Therefore, in grammatical tone languages, tone installs functional, grammatical content in a word without changing its basic, lexical semantic meaning.

The addition of grammatical function can conveniently be shown in the example of nouns in the East Cushitic language Rendille, as described in detail by Oomen (1981). In Rendille, different types of inflections affect the tonal structure of word stems. Gender inflections, for instance, can induce a change on the word stem's tone pattern such that a high tone on the penultimate syllable for masculine is deleted and, instead, a high tone is added to the ultimate syllable for feminine: *máar*, young.cow.MAS, 'male calf' vs *maár*, young.cow.FEM, 'female calf' or *ínam*, child.MAS, 'boy' vs *inám*, child.FEM, 'girl'. Akin to other grammatical tone languages, Rendille can express some inflectional features purely by means of tone changes. However, very commonly, tone languages express inflections in a combination of tonal word stem modifications and affixation. In Rendille, this is the case for number inflections. To this effect, plural is expressed by a deletion of the stem tone and the addition of the plural suffix *-ó*: *maar-ó*, young.cow.MAS-PL, 'young male cows'. Finally, grammatical tone languages can also contain inflectional affixes which do not affect the suprasegmental structure of word stems. In Rendille, gender/plural markers, which are obligatory before adjectives, are an example of non-stem-modifying inflections: *ínam-k-í ḍer*, child.MAS-MAS.SG-LINK tall, 'the tall boy'.

Note that the narrative about Rendille inflections above considerably simplifies the language's actual inflection system (cf. Oomen, 1981). In praxis, the existence of different noun classes and the interplay with phonology complicate the situation. However, the important message is that grammatical tone languages are typically diverse: Some inflectional features can be carried by the tone alone, others require an additional grammatical morpheme, such as a suffix, and yet others are expressed in non-tonal grammatical morphemes. Not all three options are available in all grammatical tone languages, defining to some degree where they should be placed on the tone language continuum. As seen for Rendille, grammatical tone languages can have inflections expressed only by tone and inflections expressed only by suffixes. This is an interesting starting point for studies of how grammatical tone is acquired compared to segmental grammatical features in natural languages.

2.1.3 Grammatical tone on nouns in Swedish

Although tone is not a feature one typically associates with Germanic languages, some Germanic languages or dialects do contain tone. One such language is Swedish. Traditionally defined as a pitch accent language (van Lancker, 1980), Swedish tones are often referred to as pitch accents, word accents, or word tones. Tones in Swedish are closely related to and applied in combination with inflectional suffixes. Thus, Swedish is arguably on the grammatical side of the lexical-grammatical tone continuum. Typical for grammatical tone languages, Swedish tones operate at word rather than syllable level. They are timed to the stressed syllables of word stems. Since understanding the Swedish tone system is important for interpreting the findings in large parts of the present dissertation, I sketch out the most important features of Swedish tones in the following.

Swedish has two distinctive tones which are traditionally referred to as accent 1 and accent 2. The tones are meaningful only in association with upcoming suffixes, although minimal pairs, only differentiated by tones, do exist. In Swedish, nominal word stems (*bil*, 'car'), for instance, carry accent 1 when they are combined with the definite singular suffix -*en*: bil_1-*en* (car-DEF.SG, 'the car'). In contrast, the same word stems are realised with accent 2 when they precedes the plural suffix -*ar*: bil_2-*ar* (car-IND.PL, 'cars'). Grammatical tones on word stems have been found to facilitate language processing as they constrain the choice of possible word endings. In fact, there is even evidence that Swedish speakers use the tones to pre-activate upcoming suffixes (Roll, 2015; Roll et al., 2015).

Swedish is a contour tone language in which tones are defined by the timing of their movements (Bruce, 1977, 2005). Although tone realisations are complexly affected by dialectal variation and intonation (Bruce, 1977; Bruce & Gårding, 1978), Swedish tones have been claimed to share one defining features: they are essentially pitch falls with different timings (Bruce, 1977, 2005). Considering Central Swedish, the dialect used for the stimuli in paper I, the low tone of the falling pitch movement (H+L*) for accent 1 is associated with the beginning of the stressed syllable's vowel. In contrast, for accent 2, the beginning of the vowel in the stressed syllable is associated with the high tone of the fall for accent 2 (H*+L). This difference in timing with respect to the stressed vowel led to the phonological description of accent 1 as a low tone and accent 2 as a high tone (Bruce, 1987; Riad, 2014). I adopt this terminology in large parts of this dissertation. An exception is paper IV, where the more low-level phonetic and psychophysiological reality of the Swedish tonal contours is discussed in the context of early processing.

While Swedish tones receive their grammatical function in combination with suffixes and are, therefore, not inherently grammatical themselves, their strong association with grammar can evoke grammatical processing (Roll, 2015; Söderström et al., 2016). In consequence, it appears as though they are indeed located on the grammatical side of tone continuum and, therefore, well-suited for a study on the acquisition of grammatical tone in language learners. They were used accordingly in paper I. The grammatical associations in the tones are further presumably strong enough to induce potential facilitation (see section 2.3) for the acquisition of new words where tone is the only cue to grammatical meaning (papers II to IV). There are likely strong neural connections between areas where tone is processed and areas that handle grammar processing in native speakers of Swedish. These connections can potentially be drawn upon in the acquisition of second language (L2) tones which – unassisted by suffixes – carry grammatical functions.

2.2 Electroencephalography and event-related potentials

An immensely useful tool in the study of how languages and, more specifically, grammatical and semantic features are processed is electroencephalography (EEG). EEG uses electrodes to measure the voltage potentials over a person's scalp. As scalp electrodes are relatively far from the neural generators, they are not sensitive to single, firing neurons (i.e., action potentials) but instead pick up summed postsynaptic potentials. A postsynaptic potential, which is evoked through electrochemical processes caused by action potentials, creates a dipole in or, more specifically, around a neuron. If dipoles are simultaneously created for large populations of neurons with similar orientations, the resulting global voltage fluctuations are measurable at the scalp (Luck, 2014). In order to use such voltage fluctuations to make valid inferences about cognitive processes, participants are exposed to a large number of stimuli (e.g., light flashes, sounds, words) which belong to two or more tightly controlled conditions which ideally differ only in the very feature that is being studied. Thus, one could potentially study how written words are processed when are meaningful or meaningless (nonsense words). All EEG responses pertaining to a certain stimulus condition or event type (here, meaningful vs meaningless words) are subsequently averaged, all timed to the onset of the target event (emergence of the word on the screen). This average is referred to as an event-related response (ERP). Finally then, comparing the ERPs of different experimental conditions, it is possible to study how, in the suggested case, the evoked neural responses to meaningful and meaningless words differ.

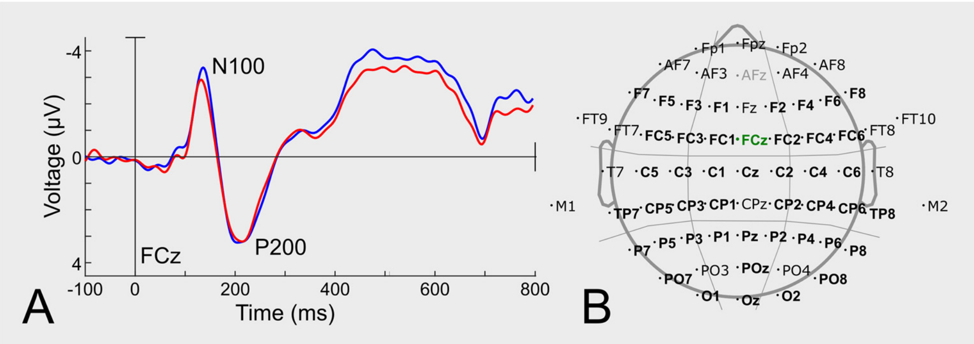


Figure 1. A. Example of ERP waveform at frontocentral electrode FCz (marked green in B), showing the early components N100 and P200 in response to relatively high (red) and low (blue) piano tones. B. Overview of all scalp electrodes used in the recordings for studies I to IV.

Amplitude differences in neural responses are often related to established ERP 'components', that is, repeatedly observed responses with characteristic latency, distribution, and deflection. Latency refers to the temporal dimension and indicates how much time has passed since the appearance of the target stimulus. Latency is described in milliseconds (ms). Distribution refers to the spatial dimension specifying at which electrodes across the scalp the effect is maximal. Deflection, finally, describes the relative difference in voltage between two studied conditions. For the well-known N400 component, for instance, which is sensitive to semantic congruence, semantically incongruent stimuli (or nonsense words) typically elicit a more negative voltage than semantically congruent stimuli (cf. Kutas & Federmeier, 2011). Voltage is measured in microvolt (µV). Most commonly, the deflection of a component mirrors the polarity of the related peak (i.e., energy maximum) of the waveform. To this effect, the increased N400 for semantically incongruent words is usually visible in the waveform as a negative-going peak.

ERP components are traditionally labelled with respect to peak polarity and latency. Hence, the N100 is a negative peak ('N') which is maximal at around 100 ms post-stimulus (see Figure 1). The P200, in contrast, is a positive peak ('P') which reaches its maximum at about 200 ms. Alternatively, the N100 and P200 are also referred to as N1 and P2, illustrating their order in the ERP waveform rather than assumed timing (Luck, 2005).

2.2.1 Grammar processing

With respect to language processing, forty years of research into mainly violation-based ERP responses have contributed to a fair but far from complete understanding of how language is processed by listeners or readers. Interestingly, in this context, is has been found that there are relatively clear differences between the processing of semantic content (e.g., Kutas & Hillyard, 1980) and that of morphosyntax (e.g., Neville et al., 1991). Although morphology and syntax are clearly distinct as they affect different linguistic levels, that is, words or sentences, they are functionally similar as both involve the combination of meaningful components into larger units. With respect to processing, there are no great differences between how listeners respond to processes of word formation (e.g., Söderström, Horne & Roll, 2016) or sentence formation (e.g., Osterhout & Holcomb, 1992; Neville et al., 1991). Thus, besides being functionally similar in a linguistic sense, morphology and syntax are also associated with similar processing (Marantz, 1997; Pulvermüller et al., 2013). Therefore, although the focus of the four studies within the present dissertation lies on inflectional morphology, I draw on examples from morphology as well as from syntax when introducing the neurophysiological processing of grammar in the following sections.

In native speakers, grammatical violations have been seen to elicit three important grammatical components: the early left anterior negativity (ELAN) at ~150 ms, the left anterior negativity (LAN) at ~400 ms, and the P600 at ~600 ms (see Figure 2).

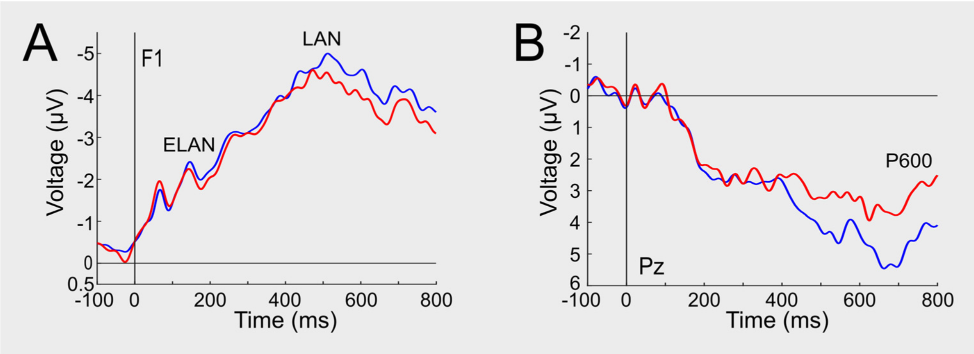


Figure 2. A. Example of ELAN and LAN component at left frontal electrode F1 in response to novel words with (blue) and without (red) grammatical content. B. Example of the P600 at central posterior electrode Pz for the same words.

2.2.1.1 ELAN

The ELAN component is a preconscious response, originally observed for word category processing (e.g., Friederici et al., 1993; Herrmann et al., 2011; Neville et al., 1991). Hence, it appears in sentence contexts where word order violations lead to the occurrence of an unexpected word category, for example, *Max's* *<u>of</u>* *proof the theorem* (Neville et al., 1991). An ELAN can, however, also be elicited for other types of automatic morphosyntactic processing such as local agreement violations (Hasting & Kotz, 2008; Shtyrov et al., 2007). The preconsciousness of the component is emphasised by the fact that it is elicited even in paradigms with distractor tasks[1] (Bakker et al., 2013; Shtyrov et al., 2007). It should be mentioned, that ELAN-type responses are not restricted to the domain of language but can also occur for other types of regularity-based combinatorial processing, such as harmony in music (typically right-lateralised, thus, early right anterior negativity, ERAN: Koelsch et al., 2000, 2002). This suggests that the type of processing observed in the ELAN component is strongly related to rule-based processing but not language-specific.

2.2.1.2 LAN

The second core component of grammar processing is the LAN. This effect is often observed in morphosyntactic contexts such as agreement violations (e.g., Osterhout

[1] To study processing outside of the scope of attention, participants often watch silent movies while passively listening to auditory stimuli. The most popular elicitation method in the context of unattended language processing is an oddball paradigm where frequent standard stimuli are infrequently interrupted by deviant stimuli. Participants typically exhibit a mismatch negativity (MMN) in response to pre-attentively perceived stimulus differences. Embedding syntactic mismatches into the oddball paradigm, the effect has been labelled syntactic MMN or sMMN in short (Shtyrov et al., 2007). It is argued to be largely identical to the ELAN, based on the same underlying process of preconscious processing of grammatical rules.

& Mobley, 1995; Roll, Gosselke et al., 2013) and is sometimes considered to be the grammar-related counterpart of the N400 (e.g., Molinaro et al., 2014). Importantly, an increased LAN is elicited for the application of regular morphosyntactic rules but not irregular ones (e.g., Newman et al., 2007; Schremm et al., 2019), and only when the rules are transparent (Rodríguez-Fornells et al., 2001). This indicates the importance of rule-based processing also for the LAN component. Further, it suggests that regular word forms are processed combinatorially while irregular word forms are not. Irregular word forms are in this context often argued to be stored as whole units while regular word forms are decomposed (cf. dual route processing, e.g., Clahsen, 1999). Interestingly, while the LAN seems strongly associated with the processing of transparent rules, it has also been f (slightly delayed) in contexts where the gender of the referent in line drawings mismatched preceding articles (Wicha et al., 2003). The line drawings are likely processed in relation to relevant semantic and grammatical cues in the preceding sentential context.

2.2.1.3 P600

The last major component of grammatical processing is the P600. The P600 is related to sentence level integration where an amplitude increase signals the need for repair or revision, typically in grammatically incongruent sentences (Kim & Osterhout, 2005; Osterhout & Holcomb, 1992). Unlike the ELAN and the LAN, the P600 seems to be less involved in rule-based processing but rather related more globally to consolidative processes. Thus, it is enhanced, for instance, for grammatical violations in both regular and irregular word forms regardless of the transparency of underlying combinatorial rules (Newman et al., 2007; Rodríguez-Fornells et al., 2001; Schremm et al., 2019). An increased P600 is also elicited for strong semantic incongruences (e.g., incongruences in thematic roles, Kim & Osterhout, 2005) where a repair or revision process is necessary to consolidate the utterance. Yet, it is very rare for semantic violations to disrupt sentence processing to a degree where utterance-level revision becomes necessary. The P600 is thus, in essence, a component that reacts specifically to grammatical violations.

Given the different conditions for their elicitation, the three main components of grammar processing can occur on their own but also together as, for instance, appears to be the case in Neville et al. (1991). In their study of word category violations (*Max's of proof the theorem*), incorrect word categories were processed as violations of morphosyntactic rules both during automatic, preconscious processing (ELAN) and at a later, likely more conscious processing stage (LAN). They further induced a need for revision and repair during consolidative processing, visible in the P600 component.

Importantly, although language-related ERP components are traditionally elicited in violation contexts, they are believed to signal general linguistic processes drawing on the same resources as canonical processing (for a functional magnetic resonance imaging study supporting this claim, see Mollica et al., 2020). This is well-attested for semantics, where amplitude changes in the N400 are observed for mildly unexpected, fully congruent words as well as for strong, violation-like incongruences or pseudowords (e.g., Kutas & Hillyard, 1980; for Swedish: Blomberg et al., 2020). The same is the case for the P600, where constructions might be unexpected and locally incongruent but fully grammatical once reanalysed (e.g., Osterhout & Holcomb, 1992). Since violation paradigms are one of the prevalent paradigms in the study of grammar processing, only very few studies have reported LAN responses outside of violation contexts (Kluender & Kutas, 1993; Krott & Lebib, 2013). These studies do, however, provide strong evidence for the general assumption that violations in native speakers merely intensify normal processing, in the same way that other resource-heavy manipulations do (e.g., Kluender & Kutas, 1993, who differentially manipulated the working memory load of sentences).

2.2.1.4 Word recognition component

Besides the three core components of grammar processing, an additional response to grammar has recently been observed: the preconscious word recognition component. At ~50 ms, the word recognition component distinguishes congruent words from pseudowords or violations. It thus assumes a preconscious linguistic gating function with respect to both semantics and grammar (cf. two MEG studies: Herrmann et al., 2011; MacGregor et al., 2012). Remarkably, well within thirty minutes of repetition, effect amplitudes to pseudowords increase, suggestive of an ongoing process of memory-trace formation for novel words in native speakers (Kimppa et al., 2015; Yue et al., 2014).

2.2.2 Tone processing

As the present dissertation aims at studying the acquisition of grammatical tone, I discuss the most important insights into tone processing in native speakers. Due to the sparsity of studies in grammatical tone contexts, I present examples from both lexical and grammatical tone languages, taking into account perceptive as well as semantic and grammatical processing.

Considering firstly the perception of tone, studies on lexical tones in Mandarin have discovered that tone perception is strongly influenced by the native status of tone. To this effect, a behavioural study illustrated that native speakers of Mandarin classify tones predominantly with respect to pitch movement rather than pitch height (Huang & Johnson, 2010). It is, therefore, argued that the shape of tone in a tonal language dictates which tonal cues native speakers attend to. Speakers of contour tone language are thought to be more sensitive to pitch movement and direction, while speakers of level tone languages language are more sensitive to pitch height. Using an MMN paradigm, a neurophysiological study was further able to show that speakers of Mandarin Chinese pre-attentively perceive even small pitch differences when those can be used to distinguish different tones. (Yu et al., 2014) For similarly small pitch differences within the same tone type, no mismatch negativity response was found. This suggests that tone perception is strongly governed not only by general tone-shape properties such as movement or height but also, more closely, by native tone types.

With respect to lexical semantic and grammatical properties of tones, processing closely mirrors what is known for lexical semantic and grammatical processing in non-tonal contexts. Tone mismatches in lexical tone contexts elicited N400 responses (Brown-Schmidt & Canseco-Gonzalez, 2004; Li et al., 2008; Malins & Joanisse, 2012). Interstingly, they typically have earlier onsets than other types of mismatches and often longer durations. This shows, not surprisingly, that tones play a crucial role for the lexical semantic content of words in lexical tone languages and that native speakers are highly sensitive to them. A similar pattern is observed in grammatical tone contexts. In Swedish, mismatches between tone and suffix (suffix effects) have been seen to elicit N400/P600 or LAN/P600 response patterns (Roll, 2015; Söderström et al., 2016). Thus, even though the tone in Swedish does not independently convey grammatical information, a mismatch with suffixes can evoke grammatical mismatch responses.

An important feature of the grammatical tones in Swedish, which are realised on word stems, is their close association with suffixes. This turns them into a cue towards upcoming structures, which native speakers consistently use during speech processing. The grammatical tones of Swedish have a strong facilitative effect on processing as their relation to suffixes allows for upcoming suffixes to be pre-activated when the associated tone appears on the word stem (tone effects). As the two tones are associated with a different number of possible endings, they have differently strong pre-activating abilities and, in comparison to each other, elicit an effect which since has been called pre-activation negativity (PrAN: Roll et al., 2017; Söderström, Horne, Frid & Roll, 2016). The impact of pre-activation can also be observed in response times measures. Correctly pre-activated suffixes have reduced response times and the more predictive accent 1 affects response times more strongly than the less predictive accent 2 (Söderström et al., 2012). It is possible, that the pre-activating function observed in Swedish extends to other grammatical tone languages, when inflections are expressed with both stem modifications and suffixes. Importantly, it is also possible, that learners of Swedish can make use of this feature to ease the acquisition process and reduce processing load.

2.3 Second language acquisition and phonological transfer

Second language acquisition (SLA) research studies how learners acquire a new language. Second languages are differentiated from first languages by convention.

First languages are commonly considered one or several languages that are acquired before a certain cut-off age, from where on onwards language acquisition is argued to be impeded and acquired languages differ from the native language(s) (Krashen et al., 1979; Meisel, 2009; Paradis, 2004; Tsimpli, 2014). The cut-off point is likely related to maturation processes in the brain but it is difficult to argue for a definite age at which language can no longer fully be acquired and processed natively. Suggestions range from toddler to early teenage years and the cut-off point is presumably relatively fluid (Paradis, 2004).

The acquisition of a second language can proceed in different settings. Most traditionally but today least commonly, a second language is acquired in target language contexts. In such conditions, the learner implicitly constructs their second language with no or limited instruction or meta-linguistic feedback (Collentine & Freed, 2005). This type of learning, called immersion, occurs, for instance, when a learners moves to a new country. Learning in immersion settings is distinct from traditional instruction-based classroom learning. The latter is more strongly focused on explicit teaching and training (Richards, 2015). There have been attempts to implement immersion as a method to language learning in traditional classrooms (Genesee, 1994). Interestingly, quickly developing technologies – from internet and social media to TV shows – have diversified language learning contexts even for primarily classroom-based learners (Richards, 2015).

While the settings mentioned above are all examples of natural language learning contexts, many studies reporting on second language learning study the acquisition of artificial languages or language material (e.g., Batterink & Neville, 2013; de Diego-Balaguer et al., 2007; Friederici et al., 2002; Havas et al., 2017). Artificial learning paradigms are used in contexts where researchers wish to have strong control over the linguistic material and/or background factors. To this end, it is possible to manipulate factors like the structure of the language input, the manner or duration of instruction, learner IQ, learner motivation, and many others. Previous studies using artificial learning paradigms have followed vastly different strategies with respect to, for instance, stimulus type and learning procedure. Stimuli range from simplified mini-versions of a natural language (e.g., mini-French: Batterink & Neville, 2013; mini-Japanese: Mueller et al., 2007; mini-Swedish: Hed et al., 2019; Schremm et al., 2017) to elaborate artificial languages (e.g., Brocanto: Friederici et al., 2002) or a selection of pseudowords (e.g., Yum et al., 2014). These can be taught in many different ways. Popular choices are pictures (e.g., Batterink & Neville, 2013; Dittinger et al., 2019; Havas et al., 2017) or strategy games (e.g., Friederici et al., 2002; Morgan-Short et al., 2010; Schremm et al., 2017). The implicit assumption is that focused learning of artificial language engages the brain's language network in the same way that classroom-based learning does, at least after an overnight consolidation period (Davis et al., 2009; Ettlinger et al., 2016).

Within the field of SLA research, a strong focus on beginner learners has been established over the past decades (cf. e.g., Gullberg & Indefrey, 2010). Studying the early stages of language learning can provide important insights into how humans are able to deconstruct new input into relevant parts and use those to make up rules and categories. Finding situations where such rapid systematisations and categorisations of incoming input are initially unsuccessful can inform about the limits of automatic language parsing but, potentially, also about what those limits are based on and how or if we can overcome them.

Previous findings in the context of initial second language acquisition have revealed some astonishing feats of the human language processing system. For instance, it was shown that listeners can detect word meaning and language formation rules simply by watching a short weather report (Gullberg et al., 2012) or that humans can extract combinatorial rules by listening to an unsegmented, continuous stream of syllables (de Diego-Balaguer et al., 2007). With respect to nominal inflections, which are at the heart of the current dissertation, a series of studies showed that just 25 or 90 minutes of learning, respectively, were sufficient for learners to achieve high accuracy rates and to evoke grammatical processing for novel inflections on novel words (Havas et al., 2015, 2017). Interestingly, phonological similarity with the native language seemed to positively influence the initial acquisition of nominal inflections (Havas et al., 2018).

2.3.1 Transfer in second language learning

An important factor which influences how second language input is processed in L2 learners is transfer. The concept of transfer was initially introduced as the way that the native language influences the perception and production of a second language (Lado, 1957). Since its original postulation, the notion of transfer has been broadened to include multidirectional, (dis)similarity-based influences amongst all languages that are part of a speakers' language inventory (e.g., Yu & Odlin, 2016). That is, there can be transfer, for instance, from the native language to a second language, from a second language to another second language (also 'third language'), or from a second language to the native language. Transfer can affect all linguistic domains, from phonology to pragmatics. In the context of the present dissertation, I focus exclusively on transfer from the native language (L1) to a targeted second language, L1-L2 transfer in short.

2.3.1.1 Positive and negative transfer

Transfer can be positive or negative (Luk & Shirai, 2009; Yu & Odlin, 2016). Positive transfer describes the facilitation of the acquisition, processing, or use of an L2 feature on the basis of similarities between the native language and the target language. Negative transfer, in contrast, can be defined as the inhibition that occurs if a target L2 feature or construction differs between the learners' native language and the target language. It has, for instance, been shown that the acquisition of number inflections on L2 nouns is delayed when nouns are not inflected for number in the learners' L1 (Charters et al., 2012; Luk & Shirai, 2009). Specifically, Japanese or Vietnamese learners experience problems in the acquisition of the English plural suffix *-s* since their native languages do not mark number on nouns. Negative transfer and related acquisition difficulties can arise for one of two reasons: Either because there is an imperfect match between the use of a feature in the L1 and the L2 or because the feature is non-existent in the learners' L1. The non-existence of an L2 feature in the L1 is sometimes referred to as neutral transfer (e.g., Selinker, 1988). In the present thesis, however, I adopt the use of negative transfer in all situations where L1 and L2 are dissimilar (Yu & Odlin, 2016).

Transfer is likely particularly important in the initial stages of second language learning, where learners have not yet formed an abstract representation of the L2 and learning is often argued to be mediated through previous language knowledge, critically including the L1 (e.g., Rast, 2010 but cf. Pienemann, 1998 for an alternative view).

2.3.1.2 Grammatical transfer

With respect to the acquisition of a second language, grammatical transfer appears to be of utmost importance (e.g., Kotz, 2009; Luk & Shirai, 2009), as illustrated in the example of plural suffixes above. Grammatical structures are differently easy to acquire depending on whether they are similar or different in native and target language. The related facilitative and inhibitory effects, respectively, can be observed both in behavioural and neurophysiological responses. Yet, they are often more evident in relatively early ERP components than in late processing or behavioural results (Andersson et al., 2019; Carrasco-Ortíz et al., 2017; Gillon Dowens et al., 2010). Behavioural attestations of positive effects of grammatical L1-L2 transfer do, however, exist, especially at early acquisitional stages (Havas et al., 2015).

In artificial learning contexts, interestingly, effects of grammatical transfer are not typically observed. Thus, learners often produce the same ERP effects regardless of whether the mismatched feature bore similarity to their L1 grammar or not (Friederici et al., 2002; Batterink et al., 2013). This might be a side effect of strong experimental focus on the manipulated grammar features. The limited number of well-controlled stimuli and rapidly increasing proficiency might cancel out the potential influence of the L1 grammar. Also, grammatical transfer is rarely specifically targeted in artificial language learning studies. One behavioural study that did make deliberate manipulations with respect to L1 and L2 grammar, did report a facilitative transfer effect based on L1-L2 similarity (Havas et al., 2015).

2.3.1.3 Phonological transfer

Besides grammatical transfer, phonological transfer also appears to be rather important in the acquisition of a new language. This is visible even in artificial acquisition settings. In a behavioural study, Havas et al. (2018) observed vast differences in the acquisition of words with native-like phonology and foreign phonology, such that unfamiliar, foreign-sounding words were significantly more difficult to acquire and had lower accuracy across a battery of tests. In a neurophysiological study with pseudowords in native speakers, Kimppa et al. (2015) found that only novel words with native phonology showed indications of memory trace build-up. Those results are suggestive of differential influences of native and foreign phonology on the processing of novel words.

2.3.2 Grammar processing in second language learners

Akin to studying processing in native speakers, neurophysiological measures can also serve as an important tool in the study of second language acquisition. They can help determine how learners process second language meaning and grammar. To this end, it has been found that L2 learners typically produce N400s for semantic processing in much the same way that native speakers do (Chen et al., 2017). Interestingly, this effect emerges as early as after 14 hours of classroom instruction or just 6 minutes of word learning in experimental settings (Dittinger et al., 2019; McLaughlin et al., 2004). The learners' processing of L2 grammar, however, tends to differ systematically from that of a native speaker, as is illustrated below.

2.3.2.1 N400 and P600

One thing that is specific for learner processing of grammar, is the fact that, during the earliest stages of L2 grammar acquisition, L2 learners produce an N400 rather than a grammar-related component (ELAN, LAN, P600) in response to grammar violations (e.g., McLaughlin et al., 2010; Morgan-Short et al., 2010). Thus, during the earliest stages of second language acquisition, grammar errors are processed as instances of incongruent, unexpected words rather than incorrect grammar.

Yet, when the learner starts to understand and internalise the grammatical rules of the L2, a first ERP component of grammar processing, the P600, emerges. This can occur quite rapidly, well within 30 weeks, in many cases within 16 weeks, of classroom learning (McLaughlin et al., 2010). While 30 or even 16 weeks might sound like a relatively long time, it is important to remember that natural L2 learners are tasked not only with the acquisition of grammar but also, simultaneously, with the categorisation and production of L2 phonology, the acquisition of a new vocabulary, etc. Given this overwhelming multitude of new information and the typically low number of instruction units per week, the emergence of a P600, signalling grammar processing, within 16 weeks is indeed rather astounding. Interestingly, this was observed even in negative transfer settings (McLaughlin et al., 2019). More strongly and solely targeting grammatical features, the P600 has been elicited in experimental settings as early as after 8 or 90 minutes (de Diego-

Balaguer et al., 2007; Havas et al., 2017). Crucially, the P600 emerged before overnight consolidation[2], that is, before memory traces could be consolidated in the brain. Such a consolidation phase has been argued to be necessary for newly acquired linguistic forms to move from the brain's memory hub, the hippocampus, to language processing areas in the cortex (Davis et al., 2009).

2.3.2.2 LAN

While, as shown above, the transition from lexicosemantic (N400) to conscious grammatical processing (P600) occurs rapidly, grammar processing at earlier latencies is not commonly found early in the acquisition process. In fact, to my knowledge, the LAN has thus far not been observed at beginner or intermediate proficiency levels. Instead, the LAN seems dependent on particularly strong entrenchment and related high L2 proficiency. It is, therefore, found, for instance, in natural L2 learners after 22 years of immersion in L2 contexts (Gillon Dowens et al., 2010).

While difficult to find in natural second language learners, the LAN has been elicited with the help of artificial learning paradigms. In focused, experimental settings, grammatical structures can get entrenched to such a degree that automatic processing, as seen in the LAN, is possible within days or months (Morgan-Short et al., 2012; Hed et al., 2019). This illustrates, that grammar processing is indeed possible after a relatively short amount of time given explicit focus and training. It likely depends on input frequency and achieved proficiency with respect to a given grammatical feature. Therefore, it is not necessarily the case that automatic grammar processing is reserved for native and near-native speakers and principally impossible to achieve for most L2 learners.

2.3.2.3 ELAN

The yet more automatic and clearly preconscious ELAN is virtually not found in natural second language learners. It is only observed if experimental conditions induce a strong, singular focus on a particular grammar structure, making the settings comparable to artificial learning (e.g., Hanna et al., 2016). In addition to this, there is, to my knowledge, one study that has found an ELAN in experimental learning conditions: Friederici et al. (2002) reported that learners trained to very high proficiency in a complex artificial language produced an ELAN after several days of intensive training. This suggests that training-induced proficiency and input frequency might be governing factors in the elicitation of preconscious grammar processing.

---

[2] Please note that the term consolidation here differs from the consolidation processes associated with the P600. While consolidation in the context of the P600 refers to sentence or utterance-based incorporation of incoming linguistic input, overnight consolidation refers to supposedly sleep-induced processes during which novel words spread to the cortex (Davis et al., 2009).

2.3.2.4 Word recognition component

Besides the above described traditional grammar components (ELAN, LAN, P600), grammar violations in the form of pseudo suffixes have also been found to elicit differences in the word recognition component in natural second language learners (Kimppa et al., 2019). To this end, advanced and beginner learners showed the same tendencies as native speakers, for instance, in the comparison of existing and novel derivations. Preconsciously recognising existing derivations, the amplitude of the word recognition component was reduced. The resulting amplitude difference between existing and novel derivations was strongest and most broadly distributed in native speakers, slightly reduced in advanced learners, and strongly reduced in beginner learners. Thus, much like for the core grammar components, the word recognition effect is most native-like in natural L2 learners at high proficiency levels.

2.3.2.5 Grammaticalisation of second language grammar

The evidently stage-based progression of grammar processing in L2 learners might be related to underlying, stage-based progressions of error processing in leaners. These are well-attested behaviourally (cf. Sorace, 1985). Learners are initially be unable to identify L2 errors. Relatively quickly, they enter a phase of error detection (step 1). With increased proficiency, they become able to correct errors (step 2) and – at high proficiency levels – they can even verbalise underlying rules (step 3). The neurophysiological results presented above are potentially related to the suggested error processing stages. When learners become sensitive to errors (step 1), they likely respond intuitively to the unexpectedness or experienced oddness of the incorrect grammatical input without being able to pinpoint the inconsistency as grammatical. This indistinct feeling of something not quite being right should elicit the observed N400 component (Foucart & Frenck-Mestre, 2012, McLaughlin et al., 2010). At the next stage (step 2), where learners are able to identify errors, they have arguably understood the grammatical features, allowing for conscious grammatical processing, as seen in the P600 component (McLaughlin et al., 2010). At the latest stage (step 3), when learners are able to verbalise rules, they have likely consciously deduced the rules of a given grammatical feature of which they have good command. At this point, the learners are aware of the combinatorial, rule-based nature of the grammatical feature and likely process it combinatorially, which may facilitate the emergence of the LAN (Gillon Dowens et al., 2010). Further internalisation and automatisation via continual exposure might eventually lead to the occurrence of an ELAN component.

The suggested progression based on observed neurophysiological findings and Sorace (1985)'s study of error processing in learners very well mirrors the notion of grammaticalisation suggested in Osterhout et al. (2008). The concept was introduced in the context of second language learners as an "instantiation of grammatical knowledge into the learner's on-line, real-time language processing system" (p. 510). As illustrated above, an L2 learner's response to grammar errors typically progresses from an initial N400 to a P600, followed by a LAN and an ELAN. This sequence of neurophysiological responses hints at an underlying development from word-based, lexicosemantic responses to a conscious, consolidative processing of newly deduced grammatical features followed by an automatic, combinatorial processing of frequent, regular L2 grammar.

2.3.3 Violation processing in second language learners

Neurophysiological studies in the context of language traditionally use violation paradigms in order to examine the processing different linguistic features. This is done based on the implicit understanding that the appearance of incongruent language would increase the processing load but not the type of processing. There are, in fact, some studies that point to the general validity of this premise. To this effect, a number of components have been manipulated to show differential amplitudes for linguistic processing outside of error contexts (N400: Kutas & Hillyard; P600: Osterhout & Holcomb, 1992; LAN: Krott & Lebib, 2013, Kluender & Kutas, 1993[3]). Similarly, a neuroimaging study has observed similar activation patterns for canonical and non-canonical language (Mollica et al., 2020). It thus appears as though the underlying assumption is correct and violation processing is presumably based on the same neural sources and processes as canonical processing and only alters the feature's processing load.

However, importantly, an increase in processing load can critically influence second language processing. Hahne and Friederici (2001) compared responses to canonical language in second language learners to responses to violated language in native speakers and found that they were equally increased. On this basis, they argue for an increased processing load in second language processing that equals that of violations processing in the native language. The claim of increased resources for L2 processing is also made based on neuroimaging findings which showed that activation in language areas was inversely proportional to L2 proficiency (Perani & Abutalebi, 2005). Assuming that processing resources are limited, the combined processing load for L2 contexts and violations might incapacitate learners in their interaction with L2 violations. As a result, L2 learners may at first be incapable of responding to violations. At the same time, they might readily process non-violated language. It is possible, therefore, that violation paradigms, while useful in native language research, might not be an ideal tool for studying L2 contexts. They might obscure the learners' true L2 abilities. The potentially problematic use of violation

[3] For a discussion on whether the word category examples of the ELAN are examples of violations or unexpected, fully grammatical language, see Steinhauer & Drury (2012).

paradigms in second learner processing studies is addressed in a controlled, artificial learning setting in papers II and III.

2.3.4 Tone processing in second language learners

An important feature to study from an L2 perspective is tone. Recent estimations suggest that there are almost 199 million second language speakers for the largest lexical tone language, Mandarin, alone, and for the largest grammatical tone language, Hausa, there are approximately 25 million non-native speakers (cf. Eberhard et al., 2020). Given these staggering numbers of second language learners, the question arises how learners cope with tone during L2 acquisition. Since there are very few studies on grammatical tone, I will primarily draw on data from lexical tone languages to illustrate what is thus far known about L2 tone perception and processing.

With respect to tone perception, it has been shown that learners at all proficiency levels struggle with correctly perceiving and categorising L2 tone. To this effect, even advanced learners have tone identification accuracies of below 30% for certain L2 tones in natural speech contexts (Yang & Chan, 2010). In single word contexts, advanced learners identify tonal words at an accuracy of below 70% with a large standard deviation (47%, Pelzl et al., 2020). Importantly, L2 tone perception is strongly influenced by the learners' or listeners' L1. Participants from non-tonal L1 backgrounds are sensitive to tone height but not to tone movement cues (e.g., Gandour, 1983; Huang & Johnson, 2010; Pelzl et al., 2019, 2020). This L1-based perceptive disadvantage impairs them not only in the identification but also in the discrimination of tones and affects their preattentive MMN responses. They respond to pitch intervals instead of functional tone differences (Burnham et al., 2015; Shen & Froud, 2019; Yu et al., 2019), demonstrating the impeding effect of negative transfer for L2 tone perception. Learners who have a tonal native language, in comparison, are able to make use of both tone height and tone direction cues in the identification of L2 tones. (e.g., Gandour, 1983; Huang & Johnson, 2010). The L1-based positive transfer enable learners from tonal backgrounds to more easily complete behavioural tone discrimination tasks and differentiate the tones preattentively (Burnham et al., 2015; Yu et al., 2019). Interestingly, Burnham et al. (2015) tested tone discrimination of learners not only from non-tonal and tonal L1s but, more specifically, from different tonal background. They found that the more similar the L1 and L2 tone, the more accurate the L2 tone discriminations, illustrating even more strongly the impact of the L1 in L2 tone processing.

There are, to my knowledge, at present only three ERP studies and two behavioural studies that target the semantic and grammatical rather than perceptive, acoustic processing of L2 tone. Two studies test English high proficiency learners of the lexical tone language Mandarin who, unlike native speakers, produce no N400 for lexical tone mismatches despite good behavioural accuracy (Pelzl et al., 2019, 2020). This likely indicates that the learners may not have lexicalised the tonal properties of words.

For grammatical tone in Swedish, in comparison, behavioural responses suggest that learners could make use of the rule-based association of grammatical tone on word stems and related suffixes. To this effect, Schremm et al. (2016) found that matched and mismatched grammatical tone elicited behavioural responses in previously tone-naïve, intermediate learners' of Swedish which resembled those of native speakers. Akin to native speakers, response times were significantly reduced when a suffix, pre-activated by the preceding tone, was correctly associated with the tone. For suffixes that were mismatched with the tone and thus not pre-activated, response times increased. This shows that intermediate learners of a grammatical tone language had likely internalised the grammatical tone system. A behavioural follow-up study further showed that training could lead to an increase in the learners' ability to actively select the implicitly pre-activated upcoming suffix (Schremm et al., 2017). Finally, a follow-up study to paper I of this dissertation tested early intermediate learners' neurophysiological processing of Swedish tones before and after intensive tone-suffix training. Before training, the learners produced no mismatch-related response. After intensive training, a LAN emerged for tone-based rule processing, as well as a pre-activation negativity for the strongly predictive tone (Hed et al., 2019). Collectively, these sparse results for the semantic and grammatical processing of L2 tone in learners suggest that grammatical tone might be easier to process than lexical tone. Learners appear to process the grammatical tone rules already at intermediate proficiency levels even when acquisition is not facilitated by positive transfer.

2.4 Research questions and hypotheses

2.4.1 Transfer in the initial acquisition of L2 grammatical tone

Papers I-IV investigate the role of phonological transfer in the processing and acquisition of L2 grammatical tone. Paper I examines how quickly linguistic processing of grammar-tone associations can occur in German beginner learners of Swedish who are expected to have negative L1-L2 transfer with respect to tone. Papers II to IV use more carefully controlled, artificial learning conditions and analyse how assumed positive and negative transfer (in Swedish and German learners, respectively) affect the processing of L2 tone and tone mismatches within the first hours of acquisition as well as after overnight consolidation. Paper II focuses on mismatch processing and paper III on learning effects (i.e., comparing taught and untaught words). Finally, paper IV examines closely what level of L1-

L2 similarity is necessary for transfer-based facilitation to take place. The underlying hypothesis is that phonological transfer has an important impact on the acquisition of L2 tone. Specifically, this is expected to become visible in the electrophysiological measures such that grammatical processing emerges more quickly in learners when there is a high degree of L1-L2 similarity, or, conversely, is delayed for learners who are unfamiliar with the grammatical function of word-level tone. Transfer may not necessarily affect late, conscious processing (e.g., P600) or behavioural responses. Finally, not all L2 tones may be identically facilitated for tonal learners, but contour tones may be easier to process than level tones.

2.4.2 Speed of L2 grammatical tone processing

Papers I, II, and III further examine the relative speed of the acquisition of grammatical processing for tonal cues in natural (paper I) and artificial (papers II and III) learning conditions in comparison to non-tonal grammatical features both in the present studies and in previous literature. The most important question in this context is how quickly a supposedly perceptually difficult, additive feature like grammatical tone can be acquired. In this context, the onset of grammatical processing of tone is studied in natural second language acquisition and comparable stages of artificial learning. The hypothesis is that despite the perceptual difficulty, training and high input frequency should facilitate a relatively rapid acquisition of both natural tone and artificial tone. In keeping with previous studies on rapid SLA, transitions from semantic to grammatical error processing (i.e., N400 to P600) are expected to emerge in beginner to intermediate learners for natural SLA or well within four hours of acquisition for artificial SLA.

2.4.3 Processing of violated and non-violated grammar

Papers II and III, finally, explore potential effects of violated and non-violated input on processing in L2 learners. As both second language material and violations have been argued to increase the processing load, we hypothesise that limited processing resources negatively affect the processing of violations in L2 contexts (paper II). Non-violated L2 input, on the other hand, is presumably equally resource demanding as violated L1 input and should therefore be more readily processed by L2 learners (paper III). This is expected to result in differential processing in the context of violations and for canonical language.

3 Methods

3.1 Participants

Two groups of each 24 participants were initially recruited at Lund University, Sweden, to participate in all of the experiments within the present dissertation. One participant from each group was later excluded from the analysis due to experienced discomfort with the recording equipment or rudimentary knowledge of a tone language (Mandarin), respectively. Consequentially, the analysed groups consisted of 23 native speakers of Swedish (10 males, mean age 24), the other of 23 native speakers of German (11 males, mean age 25). Swedish and German participants were chosen due to the close typological relation between the two languages. This would ensure that the learners had a similar linguistic background with the crucial exception of grammatical tone. While the Swedish participants came from tonal dialect areas, the German participants came from non-tonal dialect areas[4]. Besides gender and age, the groups were matched for a large number of background factors, including socioeconomic status, working memory span, number of languages spoken at above elementary proficiency, and musical training. All participants had normal hearing and could readily identify non-linguistic pitch variations and familiar variations in linguistic input (i.e., vowel length). The majority of the participants were students at Lund University, either enrolled in a local programme or studying in Sweden as part of an exchange programme. All participants were remunerated for their participation.

At the time of testing, the German participants had lived in Sweden for several months (*M* = 5.3, *SD* = 3.8, *Mdn* = 3.7) and had limited knowledge of the Swedish language. Their mean self-reported proficiency level corresponded to A1 of the Common European Framework of Reference with a range from pre-beginners (=identifying not more than single words in conversations) to early intermediates (=understanding main points of everyday conversations, ~ level B1). Grammatical tone in Swedish, which was integral to this study, is not usually explicitly taught in either first or second language teaching and, therefore, awareness of the existence

[4] German has one dialect group, Franconian, which has grammatical tone not unlike the Swedish tone system (Gussenhoven, 2004). Germans with ties to Franconian or extended stays in the related areas were not recruited. Similarly, Swedish has very few non-tonal dialect groups such as Finland-Swedish (Gårding & Lindblad, 1973). Speakers of non-tonal dialects were not recruited.

of systematic tones was low (Swedes = 15 of 23 participants, Germans = 1 of 23 participants). Further, none of the German participants and only four of the 23 Swedish participants reported awareness of the tones' ties to grammar. Yet, as tones are realised on all words in Swedish and have close ties to grammatical suffixes, we anticipated that both Swedes and Germans had statistically accumulated implicit knowledge about the grammatical tone system of Swedish from the language input (e.g., passive exposure, conversations, Swedish lessons).

For paper II, four participants – two from each group – were excluded from the analysis on grounds of poor behavioural performance on violation detection: Their response accuracy remained close to chance level (below 55%) throughout the first learning session. This exclusion did not cause any significant between-group differences regarding the background factors.

For paper IV, participants from each learner group were divided into two sub-groups based on which tones they learned as targets (as opposed to controls) in the acquisition paradigm (i.e., low/rise group vs high/fall group). The sub-groups remained matched for all background factors. Due to the uneven number of participants in each learner group ($N = 23$), however, the sub-groups were unequally large ($n = 11$ or $n = 12$). Mean imputation was employed to compensate for this in the statistical analysis.

3.2 Stimuli

3.2.1 Natural second language acquisition

To examine the processing of grammatical tone in natural SLA settings, we used 120 auditory Swedish sentences of the type 'NP got target PP':

(1)	*Kurt*	*fick*	*kopp*$_1$*-en*	*till*	*påsk*
	Curt	got	cup-DEF.SG	for	Easter
	'Curt got the cup for Easter.'				

(2)	*Kurt*	*fick*	*kopp*$_2$*-ar*	*till*	*påsk*
	Curt	got	cup-IND.PL	for	Easter
	'Curt got cups for Easter.'				

The sentences were originally recorded for Roll et al. (2015) in an anechoic chamber by a speaker of Central Swedish. Accent 1 and accent 2 were clearly distinguishable in all sentences. For every correct sentence (example (1) and (2)), a tone-mismatch counterpart (example (3) and (4)) was constructed by splicing the singular ending stem together with the plural suffix and vice versa:

(3) *Kurt fick *kopp$_2$-en till påsk*
Curt got *cup-DEF.SG for Easter
‘$^?$Curt got the cup for Easter.’

(4) *Kurt fick *kopp$_1$-ar till påsk*
Curt got *cup-IND.PL for Easter
‘$^?$Curt got cups for Easter.’

Importantly, the stimuli were created such that the target words were in unfocused, sentence-medial position in order to prevent the tones from being influenced by focus and sentence-final intonation (e.g., Bruce, 1977). To facilitate the extraction of target words from the carrier sentence for pitch manipulation, the pre-target word always ended in and the post-target word always started with an unvoiced plosive. Similarly, to ensure unproblematic splicing of target word stems and endings, the word stems also ended in unvoiced plosives. All target word stems were monosyllables.

Naturally recorded pitch was used for accent 1 and accent 2. The accents' duration and pitch movements differed slightly between words. Accent 1 (M = 137 ms, SD = 15) was overall slightly longer than accent 2 (M = 117 ms, SD = 27). With respect to tone movement, the fall of accent 1 (onset: M = 2.71 semitones, SD = 0.86, offset: M = 1.30 semitones, SD = 0.52) was distinctly shallower than the fall of accent 2 (onset: M = 6.76 semitones, SD = 1.53, offset: M = 3.08 semitones, SD = 2.25).

3.2.2 Artificial language learning

For the artificial SLA studies, we constructed monosyllabic auditory words which crucially contained a tonal element. The construction process of the words was well-controlled. To ensure that there were no differences in the quantity and quality of vowels and consonants between words, we recorded all consonants and vowels in isolation and spliced them into pseudowords. Thus, in an anechoic chamber, a male speaker of Russian produced four vowels (/a/, /ɛ/, /i/, /u/), seven initial consonants (/d/, /f/, /k/, /l/, /p/, /s/, /t/) followed by dummy vowels (/o/ or /ø/), and nine word-final consonants (/f/, /k/, /l/, /m/, /n/, /p/, /r/, /s/, /t/) preceded by the same dummy vowels. The recorded vowels and consonants thus bore phonetic features of Russian which were similar to but distinct from both German and Swedish. As a consequence, they sounded equally (un)familiar to the two learner groups. In Praat (Boersma, 2001), the recorded consonants were separated from the dummy vowels with which they were recorded. All consonants and vowels were normalised to the same power and lengthened or shortened to the same length: initial consonant = 330 ms, vowel = 350 ms, final consonant 200 ms. For plosives, silence was added as necessary consonant-initially for initial consonants and consonant-finally for final consonants. Using 10-ms transition phases, consonants and vowels were spliced into 54 monosyllabic CVC pseudowords. Eight German and three Swedish native speakers consequentially judged the words with respect to their pseudoword status as well as perceivability and producibility. Importantly, the small perception study showed that consonants and vowels had high inter-subject consistency. The d-sound in all d-initial words, for example, was transcribed and produced identically by a given participant. The 24 pseudowords whose perception was most consistent between participants formed the basis for constructing the test stimuli, see Table 1.

Table 1. Pseudowords used in the artificial learning experiment

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

Note that there were two sets of pairs in the pseudowords, the vowels /a/ and /ɛ/ always had the same 'consonant frame' (combination of initial and final consonant), as did the vowels /i/ and /u/. Thus, in the learning paradigm, it was always either words with /a/ and /ɛ/ or with /i/ and /u/ that formed the target words. The other set of words was used as control words.

Finally, four different tones were added onto each of the 24 CVC words: a high and a low level tone as well as a linearly falling and a linearly rising contour tone. The addition of the four tones to each of the 24 words resulted in a total of 96 pseudowords which were used in the learning experiment for studies II to IV. The high tone had a steady pitch of 138 Hz while the low tone had a pitch of 98 Hz. The falls and rises moved from the high to the low pitch or vice versa. Pitch was always aligned with vowel onset and pitch movements proceeded linearly over the entire duration of the vowel, starting at vowel onset. This ensured that the recognition point of tones and vowels was identical. At the same time as the vowel frequencies allowed for a vowel recognition, pitch frequencies and the presence or absence of movement allowed for tone recognition. Finally, the emergence of tone and vowel also allowed for a lexicality decision, that is, a classification of the word as a meaningless control word or as a meaningful target word. Hence, it was at vowel onset that all properties of the word could be disambiguated.

In order to teach the auditory words, we constructed visual stimuli in the form of black-and-white drawings of people with twenty-four different professions (Figure 3). The people had prominent visual features, traditionally associated with male and female, with the intention of representing the grammatical category gender (masculine and feminine) and were presented alone or in groups (two, three, or four) to represent the grammatical category number (singular and plural). The use of variable numbers in the plural condition was implemented to discourage numeric readings. Both female and male workers were of equal height and were all presented centrally on the screen to avoid excessive eye movements which would affect the EEG recordings. Finally, scrambled pictures were used for the control condition.

Figure 3. Examples of visual stimuli.
Exemplary stimuli for two of the twenty-four professions, illustrating how gender and number differences were depicted. Participants unanimously reported that visual cues for distinguishing number and gender were easy to perceive. To the very right, the gray-scale pattern is an example of the type of control stimuli used in the artificial learning study.

3.2.3 Extra-linguistic pitch perception

In order to ensure that participants perceived pitch differences equally well outside the scope of language, we created 20 consecutive piano tones ($B\flat_3$ to F_5). They were combined into 64 high-low or low-high pairs with a difference of between one and eight semitones.

3.3 Procedure

After giving their informed consent to participating in the study, participants were fitted with electrode caps containing 64 scalp electrodes (see Figure 1 for electrode placement). Two bipolar electrooculogram channels were added to monitor eye activity. During electrode application, participants filled out background questionnaires. When all electrodes had been placed and impedances had reached the required levels, experiments commenced. Participants were seated in front of a computer screen on which instructions and visual stimuli were presented. All written text, including fixation marks, was presented in white letters on a black screen, centrally aligned. Picture stimuli for the artificial learning paradigm (see Figure 2), were presented in the centre of a black screen and offset by a white background. A response box was placed on the table in front of the participants at a comfortable distance. Participants were asked to keep their index fingers on the response box at all times to ensure accurate response times. A pair of circumaural earphones was carefully put over the electrode cap for the presentation of auditory stimuli. During all experiments, EEG was recorded continuously with the exception of longer breaks, which were offered approximately every 40 minutes. During the breaks, participants were free to move around and were offered snacks and coffee. Informal comments during breaks provided information on the perceived difficulty of different elements of the experiment. Electrode cap position and impedances were readjusted after each break. The described procedure was repeated in session 2 with the exception of the informed consent form and background questionnaires.

3.3.1 Natural second language acquisition

In the actual order of experiments, the natural language acquisition experiment, which elicited data for paper I, was carried out last, that is, at the very end of session two. This was done to prevent the mismatched Swedish tones from leading to a conscious awareness of the importance of tone for the present study. An unwanted explicit focus on tone could have differentially affected the tone acquisition of papers II to IV.

To test how natural L2 learners responded to grammatical tone in Swedish and their associations with inflectional suffixes, participants were asked to listen to matched and mismatched Swedish sentence stimuli. They were tasked with hitting alternating buttons on the response box as soon as they perceived a sentence to be finished. This task was chosen to ensure that even low proficiency learners of Swedish could respond comfortably. Being potentially unfamiliar with words or the sentences' grammatical structures, learners could rely on familiarity with intonational cues in the detection of sentence boundaries. Thus, the intonation of declarative sentences in the L2 was similar to declarative sentence intonation in their L1 (Gibbon, 1998; Gårding, 1998). Choosing a delayed sentence boundary detection task that was easy to complete for low proficiency learners, we were aware that we would likely not be able to capture behavioural markers of tone validity (cf. Roll, Söderström & Horne, 2013). Native speakers served as a control group. Five test trials familiarised participants with the experimental procedure.

3.3.2 Artificial language learning

For the artificial learning part of the thesis, corresponding to papers II, III, and IV, the participants were taught the above described 24 artificial tonal words through association with meaning-assigning pictures. Each participant was taught a unique combination of words and pictures. Consonants always had lexicosemantic content (here, professions). Tone was associated with number for half of the participants and with gender for the other half. Vowels expressed the remaining grammatical category. We pseudorandomly assigned which of the tone pairs (high/fall or low/rise) was associated with which grammatical category (gender, number) and which tone within the pair was associated with which grammatical feature (feminine, masculine or singular, plural). Vowels (a/ɛ or i/u) were assigned in the same way. For an example, see Figure 4. We also pseudorandomly selected a set of six professions for each participant. Care was taken to prevent sets from containing similar, potentially confusable professions (e.g., flight attendant and waiter or flautist and violinist). The stimulus combinations for participants in the Swedish group matched the combinations for participants in the German group: thus, participant 1 in one group learnt the same word-picture pairs as participant 1 in the other group.

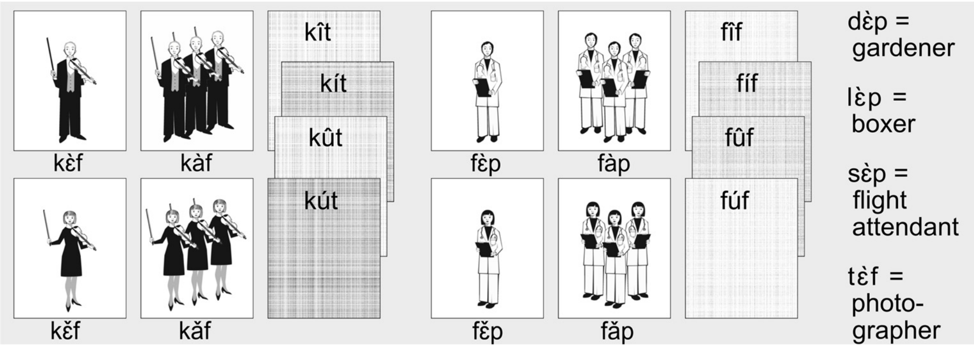


Figure 4. Exemplary distribution of the artificial words' lexical and grammatical features as associated with pictures. Two of six sets of full sets of profession stimuli and associated pictures illustrating how vowel, tones, and consonants were taught for one particular participant (participant 8) in both learner groups. The remaining four professions are listed in the male singular form. For this participant, number was expressed in the vowel (ɛ = singular, a = plural) and gender in the tone (low = male, rising = female). The vowels i and u as well as the high and falling tone were used in the control words. Tones are indicated as follows: ` = low tone, ´ = high tone, ᵛ = rising tone, ^ = falling tone.

All learners were familiar with lexicosemantic differences expressed in consonants (cf. English *sit*, *sip*, *pit*, *pip*). Additionally, both learner groups were familiar with systematic, grammatical modifications of vowels in word stems. Such modifications exist in both German and Swedish, for instance, based on the Indo-European ablaut (Swedish: *vinn!*, win.IMP, 'win!' vs *vann*, win.PST, 'won'; German: *gewinn!*, win.IMP.2SG, '*win!*' vs gewann, win.PST.3SG, 'won') or the Germanic umlaut (Swedish: *lus*, 'louse' vs *löss*, louse.PL, 'lice'; German: *Laus*, 'louse' vs *Läus-e*, louse-PL, 'lice'). Importantly, only the Swedes had experience with systematic, grammatical modifications of tones on word stems: $knut_1$*-en*, knot-DEF.SG, 'the knot' vs $knut_2$*-en*, tie-PST.PTCP.SG, 'tied' or $fall_1$*-et*, fall-DEF.SG, 'the fall' vs $fall_2$*-et*, fall-PST.PTCP.SG, 'fallen', where the tone on the word stem varies depending on the suffix. Consequently, participants were equally familiar with the meaning or function, respectively, with which consonants and vowels were associated in the present learning paradigm but differentially familiar with the grammatical role of tone.[5]

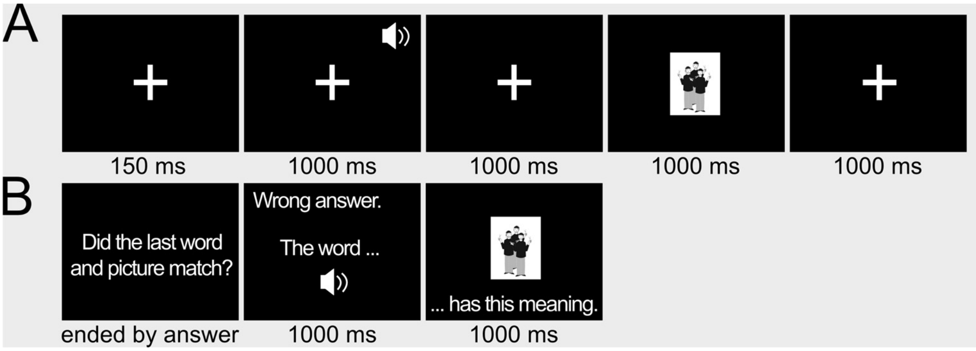


Figure 5. Procedure in the artificial learning study.
A. Slides which were part of the word-picture association paradigm. B. Slides which were added for question trials.

During learning, an auditory stimulus was presented, followed by the associated visual stimulus (cf. Figure 5). Learners were presented with thirty stimulus cycles per session. Each cycle consisted of randomly distributed repetitions of each of the 24 target stimulus pairs as well as the 24 control stimuli. After three of the 24 target stimulus pairs ('matched trials'), a question prompt appeared inquiring about the correctness of the previous word-picture pair. These three trials per cycle are referred to as 'matched question trials'. Three additional 'mismatch question trials' (or 'mismatch trials') were included in which the auditory word was presented with a non-associated visual stimulus. In these mismatch trials, words and pictures were mismatched either in number, gender, or profession. In paper II, in particular, mismatches are classified according to the phonological features on which they were based: tone or vowel mismatches (which caused errors in the grammatical domain, i.e., gender or number) or consonant mismatches (which induced lexicosemantic errors, i.e., profession). For all question trials, overt feedback was given and the auditory word was repeated together with the correct picture. Slightly varying on the basis of response times, a single cycle took approximately four minutes to complete. Thirty repetitions were chosen to ensure that overnight consolidation of the words could take place even when they had complex, combinatorial features (cf., Davis et al., 2009). Participants returned for a second learning session exactly 24 hours after the first session. The first and second session were identical in large, except for more detailed instructions and a short familiarisation phase on day 1. The instructions included information on the existence of plural and gender categories

[5] The L1 tone and the L2 tone of the artificial learning study are similar but not identical. While the L1 tone is associated with or necessarily supported by grammatical suffixes, the L2 tone is in itself grammatical.

in language, and how they would be visually illustrated in the experiment. We used drawings of one or several male and female animals (lions) for this purpose. The subsequent familiarisation phase prepared participants for the learning procedure at the example of two transparently inflected Spanish professions: *arquitect-o/a(-s)*, architect-MAS/FEM(-PL), ‘architect(s)’ and *mecánic-o/a(-s)*, mechanic-MAS/FEM(-PL), ‘mechanic(s)’.

3.3.3 Extra-linguistic pitch perception

For the background experiment on extra-linguistic pitch perception, piano tone pairs with a one to eight semitone difference between them were created. For each pair, participants were prompted to indicate whether the first or second tone of the pair was highest.

3.3.4 Electroencephalography

During all experiments, 64 Ag-AgCl electrodes, mounted in an electrode cap (EASYCAP GmbH, Herrsching, Germany) were placed on the participants’ scalp to record their neural responses to the different types of stimuli. To pick up the neural activity, 62 of the 64 scalp electrodes were used as active electrodes, one (Afz) was used as ground electrode and one (M1 = left mastoid) as online reference. The signal of the ground electrode was subtracted from all active electrodes and the reference in order to arrive at a voltage reading. The ground electrode necessarily had its own circuit in the amplifier (SynAmps^2, Compumedics Neuroscan, Victoria, Australia) to which all electrodes were connected. Ambient electrical noise picked up in the ground electrode was corrected for by subtracting the voltage of the reference electrode (reference-ground) from the voltage of each active electrode (active-ground, cf. Luck, 2005, 2014). Importantly, for all scalp electrodes, impedances were kept below 3 kΩ throughout the experimental sessions. Electrodes whose impedance increased drastically during an acquisition block were adjusted between blocks and written down in order for them to be addressed during data processing if necessary. Data was recorded with a 500 Hz sampling rate using DC mode and a standard online anti-aliasing low pass filter at 200 Hz in Curry Neuroimaging Suite 7 (Compumedics Neuroscan). The recording software automatically and invariantly set the online low-pass filter at 2/5 of the sampling rate. This was done to prevent remnants of filtered high frequencies from appearing in lower frequencies in the digitalised data, an artefact known as aliasing.

Besides the 64 scalp electrodes, a vertical and a horizontal bipolar electrooculogram (EOG) channel were recorded. In contrast to the EEG channels, the two electrodes of the bipolar EOG channels were referenced against each other rather than against M1 to create a combined vertical and a combined horizontal EOG channel. EOG channels were particularly important for noise reduction. Although scalp electrodes are targeted at detecting electrical activity produced by postsynaptic potentials in the brain, they are sensitive to all electrical stimulation. Muscle movements, which produce large spikes of electrical activity, are a potential source of noise, as muscle electricity is considerably stronger than electricity produced in the brain. We controlled for contributions by jaw, neck, and shoulder muscles by illustrating the contribution of such muscles to our participants and asking them to refrain from deliberate movement and tensing of upper body and facial muscles during the experiment. Given the length of the experiment, we selected not to ask them to refrain from blinking or moving their eyes. Instead, we used independent component analysis (ICA, Jung et al., 2000) to detect and delete eye activity from the recorded data. The EOG channels helped clearly define eye-related activity in this context. Impedances for the eye channels were kept below 10 kΩ. Once recorded and digitalised, the data was re-referenced to average reference. Expecting lateralised effects (LAN, ELAN), average reference was chosen over the more traditional combined mastoid references which has been argued to potentially dilute the significance of lateral effects (Dien, 1998). We subsequently applied an offline high-pass filter of 0.01 Hz to reduce slow drifts in the voltage. A low-pass filter at 30 Hz was further applied to reduce the impact of unwanted electricity on the findings such as potential power supply noise (AC frequency = 50 Hz). Thus filtering the data should allow for the emergence of typical cognitive components (Luck, 2014; Tanner et al., 2015). After filtering, we extracted epochs of 1200-ms length including a 200-ms baseline for all analysed stimuli, that is, pictures and words in the artificial learning paradigm as well as tone onset and suffix onset in the natural SLA settings, and finally tone onset in the piano tone part of the study. Collectively, we submitted all trials to an independent component analysis and removed components which were clearly related to eye activity or which were unambiguously related to bad channels due to loss of contact. This was checked against the data acquisition notes. All epochs which at this point still exceeded ±100 µV and thus likely contained unwanted noise were excluded from further analysis.

3.4 Data analysis

3.4.1 Behavioural data

The behavioural measures that were recorded were Response Times (RT) for all experiments and Response Accuracy for the main experiment of papers II to IV and for the piano tone experiment in paper I. For RTs, we either submitted averaged raw data to statistical analyses (paper I) or applied log transformations prior to statistical analyses (papers II-IV). For Response Accuracy, we similarly used averaged raw accuracy data (papers I & III) or d' scores (papers II & IV). Behavioural data were analysed with the help of a mixed Analysis of Variance (ANOVA) with relevant experimental factors (e.g., 'Word Type' [target vs control words] or 'Mismatch Type' [vowel mismatch vs tone mismatch vs consonant mismatch]), temporal factors (e.g., 'Session' [session 1 vs session 2]), and between-subject factors (e.g., 'Learner Group' [Swedes vs Germans]).

3.4.2 Electrophysiological data

For the analysis of the electrophysiological data, time windows were constructed with respect to sliding time windows (paper I), based on prominent peaks in neural activity (papers II & III) (global root mean square, gRMS, Lehmann & Skrandies, 1980), or with reference to previously observed effects (paper IV). For papers I and III, the electrophysiological data was further arranged into nine pre-defined topographical regions (left, mid, right by anterior, central, posterior) with five electrodes per region (bold electrodes in Figure 1). Mixed ANOVAs were consequentially carried out with the corresponding topographical factors as well as relevant experimental, temporal, and between-subject factors. For papers II and IV, relevant electrode clusters were established through permutation analyses. Mixed ANOVAs were carried out for paper IV for any observed clusters with relevant experimental, temporal, and between-subject factors.

The learner groups were analysed separately and effects compared with the help of two-tailed independent samples t-tests in paper III. Wherever we deemed this would lead to additional insights, correlation analyses were conducted between electrophysiological data and behavioural data (i.e., RTs, Response Accuracy, or background factors such as Level of Swedish for the German participants).

4 Results

4.1 Paper I

Paper I investigated the acquisition of a natural grammatical tone language, Swedish, by a group of German participants. While typologically similar to Swedish, German critically differs from the target language with respect to tone. The lack of word-level tone in German was hypothesised to induce negative transfer conditions for the studied feature. The study was based on the suggested ability of learners to rapidly detect rules and regularities of frequently presented cues in a new language (Gullberg et al., 2012; Havas et al., 2015; 2017). For classroom-based learning, this has been seen to evoke a semantic N400 after just 14 hours of tuition (McLaughlin et al., 2004) and a grammatical P600 within 20 weeks (McLaughlin et al., 2010). We wanted to expand on this and investigated whether even non-explicitly taught but frequent tone-grammar patterns could elicit native-like processing after moderate periods of second language acquisition. While tone is generally considered a difficult L2 feature to master, Swedish tone is abundant in the input and bound by a finite set of morphosyntactic rules which makes it a highly reliable cue for native speakers. We assumed that it should be possible for learners, despite previous unfamiliarity with word-level tone, to quickly deduce the frequent, systematic Swedish tone rules and internalise them. Previous behavioural data (Schremm et al., 2016; 2017) from intermediate learners with non-tonal language backgrounds strengthened this assumption: Their response times to grammatical tones in Swedish were indicative of rule-based, predictive processing. We thus expected learners at low proficiency levels to have noticed the tones' systematic role in language and, potentially, to have internalised parts of the underlying rule system. Therefore, we anticipated that the learners would exhibit previously observed responses related to processing of grammatical error in learners (i.e., N400 or P600) and resemble native speakers in their processing of Swedish grammatical tones even at early stages of the acquisition process. We included a group of native speakers to compare the learners' results against.

The results of paper I showed all expected effects for the native speakers: a pre-activation negativity for the differentially strongly pre-activating tones (tone effect) as well as an N400 and P600 for tone-suffix mismatch (suffix effect). This closely resembled previously found native speaker results. In the German learners of Swedish, on the other hand, the Swedish tone-suffix combinations did not evoke any of the native speaker effects related to rule-based tone processing. Instead, the learners produced a difference potential for the tones at 400-600 ms: high tones (i.e., tones with a steep fall on the stressed vowel) had a reduced negativity compared to low tones (i.e., tones with a low, shallow fall on the stressed vowel). A very similar difference potential was observed in the comparison of high and low piano tones. We further found an expansion of the effect for the grammatical tones from central to right-lateral electrodes which was positively correlated with Swedish proficiency (see Figure 6).

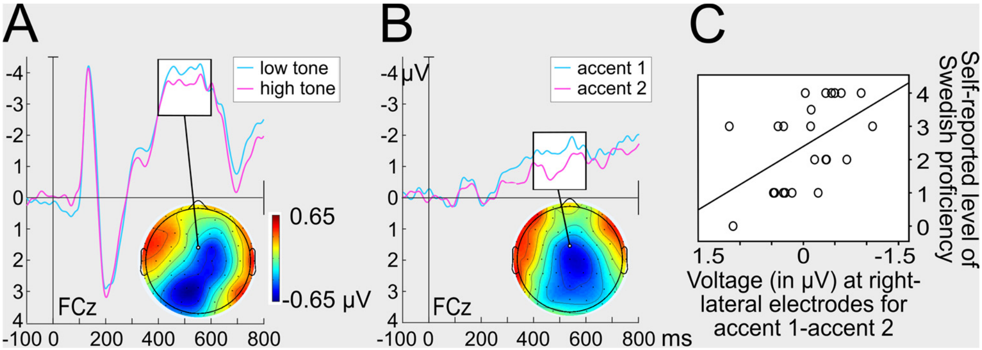


Figure 6. Beginning learners' processing of piano tones and Swedish grammatical tone.
A. ERP activity at frontocentral electrode FCz for piano tone pairs in the German group. Subtraction topography for low tones minus high tones for the analysed time window (400-600 ms). B. ERP activity at FCz for Swedish tones in the non-advanced German learners. Subtraction topography for accent 1 (low-onset tone) minus accent 2 (high-onset tone) for the analysed time window (400-600 ms).C. Correlation of the learners' Swedish listening proficiency and subtraction amplitude of accent 1 and accent 2 at right-lateral electrodes.

4.2 Paper II

Paper II investigated the effect of positive and negative phonological transfer on beginner learners' behavioural and neurophysiological responses to mismatches. Previous research has suggested that the native language processing system plays an important role during the initial stages of L2 learning (Rast, 2010). During very first contact with an L2, learners necessarily have to process L2 input with help of previously acquired phonological and morphological rules. This can lead to increased acquisition accuracy for learners who are familiar with the L2 sounds and rules (Havas et al., 2015, 2017). The underlying facilitation based on L1-L2 similarity can also manifest as facilitated processing, even when it is not measurable in behavioural responses (Andersson et al., 2019). Paper II added to previous findings by comparing the potential influence of facilitation on behavioural and neurophysiological mismatch detection responses during the initial acquisition of words with grammatical tone.

For this purpose, we used monosyllabic nouns with inflections expressed by means of grammatical tone and vowel modifications taught through picture referents. Swedish and German learners were expected to show grammatical processing of mismatches well within four hours of training. We further hypothesised that the Swedish participants' behavioural and neurophysiological responses would be affected by positive transfer based on their familiarity with grammatical tone. They would likely produce a P600 for tone- and vowel-based grammar mismatches and an N400 for consonant-based lexical mismatches. For the Germans, whose acquisition we assumed to be impeded due to unfamiliarity with the words' tonal feature, we anticipated slow response times and low response accuracy in the detection of tone mismatches as well as a delayed onset of grammar-related ERP responses. The German learners' responses of the vowel- and consonant-based mismatches, on the other hand, would likely not be affected by the tonal component and thus similar to those of the Swedish learners.

The results for paper II showed no significant differences in behavioural results between learner groups. Both groups had a comparably low accuracy for tone mismatches and their RTs were longer for this mismatch type. In the neurophysiological data, however, a facilitative effect of phonological transfer became apparent. The Swedish learners produced an N400 in all mismatch conditions, while the Germans only exhibited differential N400 responses for vowel- and consonant-based mismatches compared to matched word-picture pairs. Contrary to predictions, we found no components of grammatical processing in either learner group. The N400 emerged regardless of whether the mismatch in the pictures was grammatical (number or gender) or lexicosemantic (profession), see Figure 7. Besides the N400, we observed a posterior N1 component which was delayed in the Swedes compared to the Germans.

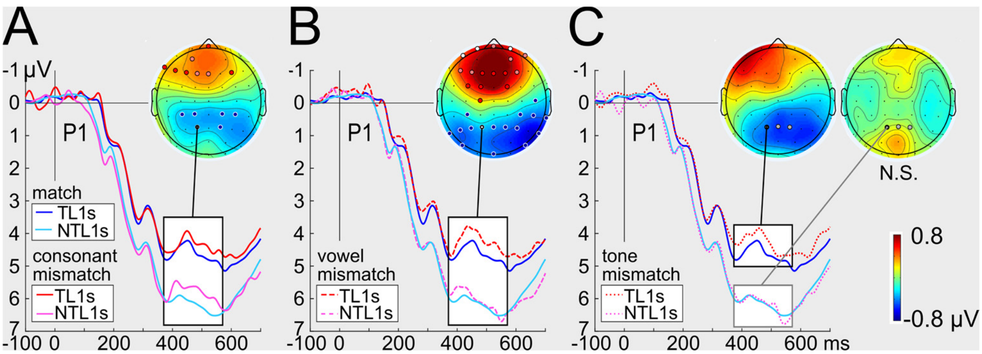


Figure 7. Responses to mismatches during four hours of word-picture association learning.
ERP activity at left-posterior electrode P1 and subtraction topography for match minus mismatch during the effect time window (370-570 ms). Electrodes involved in significant clusters are highlighted. A. N400 for mismatches with the words' consonant feature (i.e., profession). B. N400 for mismatches with the words' vowel (i.e., grammatical category: gender or number). C. N400 for mismatches with the words' tone (i.e., grammatical category: gender or number) for Swedes (TL1s, left) and Germans (NTL1s, right)

4.3 Paper III

Paper III further studied learners' ERP responses during the acquisition of artificial grammatical tone. However, instead of looking at error processing, the study examined learning effects, specifically, the congruent processing of meaningful, novel tonal words in comparison to repeated but untaught control words. We targeted the processing of violation-free linguistic stimuli on the basis of the assumption that the occurrence of both semantic and grammatical components is not limited to error contexts but instead indicative of general linguistic processes. For semantic processing and the N400, this assumption is widely accepted, presumably strengthened by the fact that the N400 was initially discovered in the context of moderate incongruence rather than violation (Kutas & Hillyard, 1980). Grammatical components, on the other hand, most commonly emerge in violation contexts. However, recent studies (e.g., Krott & Lebib, 2013) have emphasised the view that ELAN, LAN, and P600 are not violation-dependent and elicited as part of general grammatical processing. Thus, we should observe grammatical components in the comparison of stimuli which differed only with respect to whether or not they were associated with a grammatical function. In line with previous results on the rapid acquisition of new words (e.g., Havas et al., 2017), we expected the processing of novel grammatical words in canonical contexts to proceed rapidly when facilitated by L1-L2 similarity. For learners who were unfamiliar with one of the words' phonological features (i.e., word-level tone), we anticipated processing of the grammatical L2 words to be delayed.

The results for paper III were line with our hypothesis: outside of violation contexts, novel tonal words with grammatical function elicited an ELAN, LAN, and P600 within 20 minutes of association learning for Swedish learners, see Figure 8. We also observed session-initial differences between the learnt words and untaught control words in the early word recognition component. For a group of German learners with an assumed negative transfer background, the tonal words also evoked a P600, indicative of grammar processing, within 20 minutes, and a LAN on the second day of acquisition. A correlation analysis suggested that German learners with good acquisition accuracy displayed LAN-like tendencies already on day 1. No preconscious processing, that is, differences in ELAN or word recognition component, were found in the German group. A correlation with extra-linguistic tone perception skills, however, suggested a tendency for ELAN-like processing in negative transfer conditions if learners had good pitch perception and discrimination abilities.

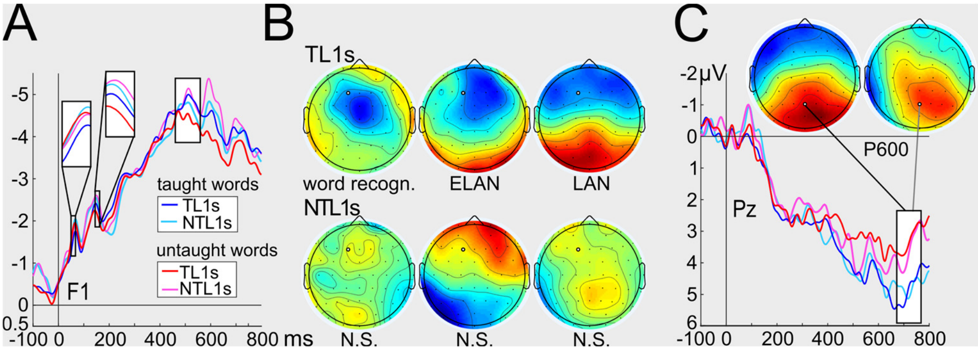


Figure 8. Responses to taught words during the first twenty minutes of word-picture association learning.
A. ERP activity at left frontal electrode F1 (marked in B) for Swedes (TL1s) and Germans (NTL1s). Three analysed time windows are highlighted. B. Subtraction topographies for taught minus untaught words for the three earliest effects in the Swedish (above) and German (below) learner group. C. ERP activity at posterior electrode (Pz) showing the P600 effect in both participant groups. Subtraction topographies show the results for Swedes (left) and Germans (right)

4.4 Paper IV

Paper IV, building up on the previous papers, investigated which degree of formal and functional similarity is necessary for transfer-based facilitation of linguistic L2 processing to occur. While previous studies on tone acquisition and processing has typically focused on the perception and perceptive processing of tone, paper IV examined the linguistic processing of tone. The word recognition component and the N400/LAN time window were of particular interest in this context, as effects of L1-L2 transfer have previously been observed at both latencies. The word recognition component, for instance, has showed evidence of memory trace formation only for pseudowords with L1-like but not foreign phonology (Kimppa et al., 2015). The LAN, on the other hand, has been seen to emerge faster for L1-like grammatical constructions (agreement, Gillon Dowens et al., 2010) while the N400 was reduced for L1-L2 cognates compared to non-cognates (Midgley et al., 2011). Importantly, the LAN/N400 time window also responded differently to contrastive tones in linguistic as well as extralinguistic contexts, which makes it particularly interesting in the context of tone processing (Gosselke Berthelsen et al., 2018; Kochančikaitė et al., 2020).

The results for paper IV revealed that the word recognition component was reduced compared to all other conditions only for target words with a falling pitch contour in tonal learners, see Figure 9. This likely signals complete memory trace formation for words with the highest level of formal similarity. No difference in the early word recognition effect emerged for tonal learners who acquired words with low and rising contours or for non-tonal learners. The late component in the LAN/N400 range, in comparison, was not affected by transfer. It was globally increased for target compared to control words, for level compared to contour tones and for trials on day 1 compared to trials on day 2.

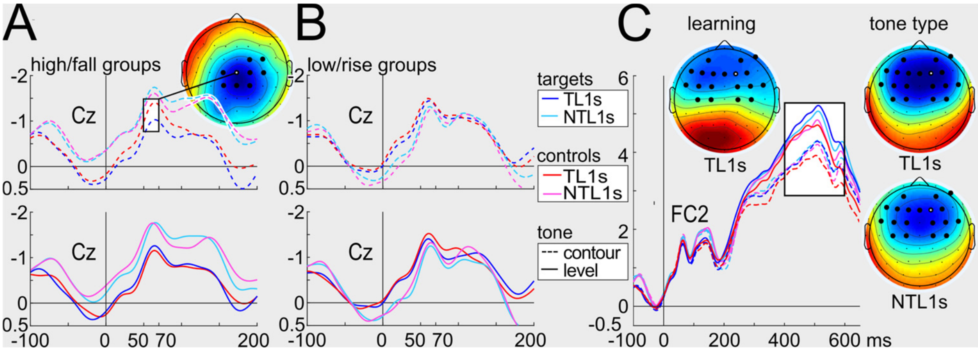


Figure 9. Effects of tone similarity and prominence during four hours of word-picture association learning.
A. ERP activity at central electrode Cz for the high/fall subgroups of the Swedish (TL1) and German (NTL1) participants. Contour tone shown above, level tone below. Subtraction topography for taught contour targets (falls) minus untaught contour controls (rises) in the significant time window (50-70 ms) for the Swedish high/fall participants. Significant cluster electrodes highlighted. B. ERP activity at central electrode Cz for the low/rise subgroups of the Swedes and Germans. Contour tone shown above, level tone below. C. ERP activity at right frontal electrode FC2 and subtraction topographies for taught minus untaught words (left) and for level minus contour tones (right). Swedes above, Germans below. Significant cluster electrodes highlighted.

5 Discussion

5.1 Transfer in the initial acquisition of L2 grammatical tone

One of the key topics of the present dissertation is the differential influence of negative and positive transfer on the acquisition and processing of L2 grammatical tone. We examined assumed negative transfer by looking at possible inhibiting effects on the acquisition of L2 tone in a natural language based on unfamiliarity with word-level tone. Paper I studied negative transfer conditions in the context of the natural language Swedish and papers II and III investigated this in the context of artificial tonal words. Anticipated facilitative effects of positive transfer, which might occur due to L1-based familiarity with word-level grammatical tone, were examined in papers II, III, and IV.

5.1.1 Negative transfer

Paper I demonstrated that beginner learners with a non-tonal L1 were unable to acquire frequent but non-explicitly taught tone-suffix rules in the context of a naturally acquired second language. We thus neither observed expected evidence of predictive processes at tone onset nor indications of mismatch responses for tone-suffix mismatches. Instead, the learners appeared to react to the tone-initial pitch levels in much the same way that they responded to musical pitch[6]. We did, however, find a lateral expansion of their ERP response to grammatical tone. We interpreted this effect as a sensitisation to pitch in the context of language, which might be an important precursor for tone processing.

In natural SLA, beginner learners who are not familiar with word-level tone from their native language could not deconstruct and acquire the L2 tone as quickly as originally assumed. Schremm et al. (2016) had found native-like behavioural responses for mismatched grammatical tone in Swedish already in intermediate learners, which prompted us to expect that tone acquisition effects would be visible already in beginner learners. The current findings suggest that the threshold for deducing the tones' involvement in language and the resulting language-related

[6] This response has since been replicated in Kochančikaitė et al., (2020).

processing must lie at intermediate proficiency. Until then, learners who are unfamiliar with word-level tone presumably do not detect its linguistic relevance. This illustrates the difficulty of the acquisition of L2 tone-grammar relations in learners who are unfamiliar with word-level tone and who are thus suggested to have negative transfer conditions with respect to L2 tone. Interestingly, it has since been found that tone-suffix training can elicit tone processing in learners at the beginner/intermediate threshold even when there was no positive L1-L2 transfer (Hed et al., 2019).

Paper II more carefully targeted how grammatical tone could be acquired in learners who were not familiar with word-level tone from their native. The same group of German learners that participated in the study for paper I was taught artificial grammatical tone. We observed that mismatches in the picture referents based on the words' tone were significantly more inaccurately and slowly detected than mismatches based on familiar phonological features, that is, vowel and consonant. Similarly to findings in a previous study (Havas et al., 2018), this suggests that behavioural responses for mismatch detection in L2 grammar are considerably delayed when the phonological L1-L2 similarity is low. We further found no evidence of mismatch processing of tone-based mismatches. Yet, the German learners did produce a mismatch response (N400) for mismatches in the picture referents when the mismatches were based on familiar phonological features. Thus, only features which closely resembled the learners' native language in form and function appear to be able to elicit mismatch-related responses in initial stages of learning. Together, these responses to mismatches in the context of artificial learning further substantiated the general idea of potentially impeded acquisition of tone for learners with a non-tonal native language and emphasise the importance of L1-L2 similarity in initial L2 acquisition (e.g., Rast, 2010).

Paper III, finally, operated on the assumption that mismatch processing and the processing of fully congruent words share common grounds and both should elicit grammatical ERP responses (e.g., Krott & Lebib, 2013; Mollica et al., 2020). On this premise, paper III investigated learning effects. As had been seen in a previous study in an L1 context, no changes in the word recognition component emerged in the taught words, presumably because they had unfamiliar phonological properties (Kimppa et al., 2015). However, Germans produced responses pertaining to the conscious processing of the taught words' grammatical features within just 20 minutes of acquisition. While it is unclear whether the grammatical processing was related to the functional content of the tones or the vowels, or both, it is still interesting that grammatical processing emerged in the context of tonal words. The added, unfamiliar tonal component of the words did not stop learners from acquiring them. In fact, the mismatch detection results suggested that even the tonal rules themselves are to a large degree acquired by most learners. Hence, learners with negative transfer with respect to word-level tone can process words with grammatical tone rapidly and the tones do not disrupt their acquisition. Interestingly, learning accuracy and pitch perception abilities appeared to enable the emergence of more automatic processing in individual learners. This suggested that overall, negative transfer might inhibit such more automatic processing learning-initially but that inhibitory effects can be overcome to some degree by non-linguistic, domain general abilities and factors such as learning aptitude or motivation or auditory acuity.

Summarising the processing and acquisition of grammatical tone in a negative transfer group, unfamiliarity with phonological features may negatively affect the acquisition of L2 morphosyntax such that the tonal feature is not acquired as rapidly as might have been expected. This became transparent, in particular, in the example of mismatch processing in both natural and artificial contexts. Mismatch-related components appeared only for features where L1-L2 familiarity was strong. The negative transfer group further showed no evidence of automatic processing of the grammatical novel words. Yet, some learners within this group, displaying different facilitative abilities, seemed to be able to overcome the inhibitory effect that the negative transfer conditions seemed to impose on them. Thus, learners with a low degree of L1-L2 phonological familiarity for a given feature are likely initially impeded in its acquisition but a number of not necessarily language-specific factors might compensate for the lack of L1-based familiarity in L2 acquisition.

5.1.2 Positive transfer

Paper II targeted the acquisition of grammatical tone in learners who were familiar with the feature from their native language. We expected these learners to show facilitated learning effects suggestive of positive transfer, although there was no full functional overlap between the L1 tone and the L2 tone.

Paper II revealed no apparent facilitation effect within the behavioural results. On the contrary, tone mismatches were significantly less accurately and less quickly detected than vowel and consonant mismatches. Tone mismatch detection was slightly more accurate and faster in the Swedes than the Germans but the differences were not significant. With regards to mismatch processing, on the other hand, paper II found that tone mismatches elicited the same mismatch response as vowel and consonant mismatches (N400). Equal familiarity with the L2 tones, vowels, and consonants and their respective functions allowed for the detection of mismatches with all three features in the Swedish learner group. This is particularly interesting in the context of the German learners who did not produce a mismatch response in the unfamiliar tonal condition. Hence, as has been seen previously, transfer facilitation might affect the processing of L2 input but not so strongly that it influences mismatch detection behaviour (Andersson et al., 2020). The elicitation of an N400 rather than the N400 rather than the expected P600 might be related to the fact that L2 tones are difficult perceptually even in presumably facilitated contexts. This is supported by behavioural findings both in the present thesis and in previous literature. To this effect, Lee et al. (1996) found that while tonal L1 listeners were better than non-tonal L1 listeners at identifying and discriminating L2 tones than speakers with a non-tonal L1, they still performed considerably worse than native speakers (Lee et al., 1996). The general perception difficulty associated with tone may also explain the discrepancy between tones and vowels in the behavioural results in the Swedish learner group. Alternatively or additionally, the occurrence of the N400 could also be explained by the implementation of the mismatch in the picture referent or to the use of a mismatch paradigm (see section 5.3).

Paper III observed a full range of conscious and even preconscious grammatical components in the processing of the acquired grammatical words. As mentioned before, it is impossible to determine whether the tone-grammar associations, the vowel-grammar associations or a combination thereof was the driving force behind the grammatical ERP components in the non-mismatch context. Yet, the extremely rapid elicitation of a full range of grammatical responses was unexpected. In comparison to the findings in the negative transfer group, it becomes apparent that the preconscious processing might be emerge in relation to and be indicative of positive phonological transfer. New words can presumably latch onto pre-existing, fine-tuned tone-grammar-associated structures in the brain. This is supported by previous research which suggests that L1 and L2 processing is carried out by the same neural networks (Perani & Abutalebi, 2005). Being able to use the brain's entrenched neural pathways for the L1, preconscious processing becomes possible.

Paper IV, finally, more closely investigated different degrees of L1-L2 similarity. It examined the role of different tone types in the processing and acquisition of L2 tone and suggested that the word recognition component was very narrowly affected by L1-L2 similarity. As such, it was reduced only for learnt words with falling tones in Swedes. This suggested that rapid memory trace formation could only take place words which carried the tone which has been argued to be functionally most important in the Swedish tone system (Bruce, 1977, 1987, 2005). This finding was not anticipated. Previous studies had typically shown a more general transfer-related facilitation effect, indicating that speakers of contour tone languages should excel at perceiving 'any' tone movement and directionality cues (Burnham et al., 2015; Gandour, 1983; Huang & Johnson, 2010; Shen & Froud, 2019; Yu et al., 2019). Consequently, the expectation was that contour tones were the main facilitating factor in the tonal L2 words. Instead, the word recognition component only showed a facilitation effect for functionally important falling tones. This agrees with findings for tonal pseudowords whose processing was strongly influenced by the potential meaningfulness of tones with respect to L1 knowledge (Braun & Johnson, 2011). Yet, while narrowly motivated response patters suggested a strong influence of phonological transfer in the word recognition component, the later anterior negativity component proved relatively unaffected by transfer effects. Instead, ERP amplitudes of the (L)AN for all learners were considerably reduced to words with contour tones compared to words with level tones. The (L)AN similarly experienced a reduction after consolidation as well as for untaught compared to taught words.

Summarising the findings of the suggested positive transfer group in the present dissertation, it appears as though that an L1-based familiarity with phonological features may positively affect the acquisition of L2 morphosyntax. Especially the emergence of preconscious processing seems to depend on familiarity with the L2's phonology. But also mismatch detection responses were throughout the dissertation only elicited for familiar features. Strikingly, while familiarity with tone generally seems to facilitate L2 tone processing, rapid word trace formation and related preconscious processing only seem to be possible when there is a very strong degree of functional and formal similarity. Thus, strong phonological L1-L2 similarity and positive transfer can generally be argued to facilitate L2 processing while different degrees of similarity might be necessary for the emergence of different components of language processing.

5.1.3 Transfer effects at different processing stages

Within the findings for German and Swedish learners, the present dissertation demonstrated the importance of phonological L1-L2 similarity for morphosyntactic processing in the initial stages of L2 acquisition. It revealed differential effects of negative and positive transfer but also illustrated that transfer-effects can be overcome at certain processing levels and that successful acquisition can occur regardless of the underlying processing. Interestingly, the observed transfer-related results suggest that the earlier and more automatic the ERP response, the more crucially it depends on L1-L2 similarity learning-initially. Only those phonetic features (here falling tones) that have a direct functional relevance in the learners' L1 can affect ultra-rapid ERP responses.

A similar claim towards the importance of L1-L2 similarity for early, preconscious processing has previously been made with respect to the acquisition of new words in native speakers (Kimppa et al., 2015). We observed that a strong degree of formal and functional similarity was also necessary for rapid memory trace formation and, consequentially, preconscious word recognition in second language learners. The ELAN response, in comparison, could occur for all Swedish learners. It was thus seemingly facilitated by all tones, not only those with falling contour tones. In consequence, global L1-familiarity with the L2 phonological concept rather than its concrete phonetic manifestations might be sufficient to elicit ELAN responses in contexts with relatively low processing costs. The role of phonological transfer for the emergence of ELAN components in the context of second language learners has, to my knowledge, not yet been targeted. Second language learners rarely exhibit ELAN effects (e.g., Mueller et al., 2005) which potentially complicates systematic studies of transfer. The present thesis is just first to suggest a systematic involvement of phonological transfer in the elicitation of the ELAN. Importantly, both the word recognition component and the ELAN could only appear outside of violation contexts. This is probably related to the availability of more processing resources and should be studied in more detail (cf. section 5.3).

At later processing stages, processing appears less affected by phonological transfer. For the emergence of the LAN, L1-L2 similarity is likely facilitative in the very beginning. Yet, training, consolidation, and increased proficiency can lead to the emergence of LAN effects even when the targeted L2 feature is not present in the L1. For syntactic transfer in a natural language, Andersson et al. (2019) found a LAN-like effect only for learners who were familiar with the studied L2 feature. With increased proficiency, however, it is likely that the observed transfer effect would reside and a LAN might emerge even in learners who are initially affected by negative transfer. This claim is emphasised by the emergence of a LAN in non-advanced language learners after focused training (Hed et al., 2019). Further, also in agreement with Andersson et al. (2019), the findings in this dissertation suggest that the P600 and behavioural responses are largely unaffected by L1-L2 similarity and transfer.

In summary, transfer predominantly affects early, automatic processing. Close formal and functional L1-L2 similarity is necessary for ultra rapid L2 processing, while a weaker degree of similarity is required for facilitation of slightly later, yet still preconscious processing (ELAN). Even later, more conscious processing (LAN) may still initially be affected by transfer but the effect decreases at increased L2 proficiency levels (e.g., Gillon Dowens et al., 2010). Late, conscious processing, as seen in the P600, is apparently unaffected by L1-L2 similarity, as are behavioural grammaticality assessments. The facilitation effect can likely be argued to operate through entrenched native language networks (Perani & Abutalebi, 2005), formed in the context of grammatical tone by strong phonology-grammar associations in the L1. The faster and more automatic the effect, the stronger the dependence on strong and direct connections within the L1 neural network for L2 input to be processed through it.

5.1.4 Limitations

There are some limitations to the interpretation of the results above. Firstly, the discussion of the natural SLA findings was based exclusively on a group with low L1-L2 similarity with the studied feature. I argue that the evidence of unsuccessful acquisition of the tone system is based on negative transfer. A claim that is difficult to substantiate without control conditions. The addition of either a familiar feature or control groups with high L1-L2 familiarity would more convincingly have illustrated the suggested impact of phonological transfer on the acquisition of

Swedish grammatical tone. The inclusion of control conditions might also have revealed other possibly difficult factors in the acquisition of Swedish tones, such as the potential complexity of the Swedish tone-suffix system or the implicitness of their acquisition. Unaware of the relative influence of those factors, definite claims about the relative impact of negative transfer in this context are difficult.

Maybe not a limitation per se but surely important to mention again is also that the Swedish L1 tone system and the taught artificial L2 tones were similar but not identical in their grammatical function. In Swedish, tones are assumed to support but not carry grammatical function. In the artificial words, the tones themselves bore grammatical functions. Yet, both tone types were related to inflections and it is likely that there are strong connections between tone and grammar processing in native speakers of Swedish which we hypothesised the L2 grammatical tone processing could facilitate from. This was indeed observed in the results suggesting that the grammatical function of the tones in Swedish was indeed sufficiently strong to facilitate the acquisition of a new, slightly different grammatical tone system.

A final limitation relates to the emergence of grammatical components (ELAN, LAN, P600) for words with grammatical tone. I discuss them with respect to differential transfer conditions based on the presence of word-level tone in the learners L1. Yet, it is unclear whether the observed rule-based processing of the words is in fact based on the grammatical features related to the tonal component. The grammatical features related to the word's vowel or even an interplay of the two (Swedish tone is crucially timed to stressed vowels) might have affected the results. That is to say, although the results differed systematically in ways that we would expect to be caused by transfer of phonology-grammar relations, it is in principle possible that the Swedes' rapid, full-fledged grammatical responses were based on the grammatical function of the vowels rather than the tones. Importantly, however, the vowels of the artificial words were strongly controlled to have equal similarity with both L1s with respect to phonetic and functional properties. A differential processing of the vowels, while possible, would be unlikely to cause the observed patterns of processing differences between Swedes and Germans. This argument is strengthened by very similar processing of vowel-based but not tone-based mismatches in the two groups.

5.2 Speed of L2 grammatical tone processing

5.2.1 Natural second language acquisition

Paper I examined the speed at which natural second language learners can acquire and process L2 grammatical tone as compared to other grammatical features. Previous studies suggested that four to eight weeks of classroom tuition were sufficient for a mismatch-related N400 to emerge in grammar violation contexts and that, within 16 weeks, even the P600 component emerges (McLaughlin et al., 2010). For Swedish tone, a behavioural study suggested that learners at intermediate proficiency were able to process tone linguistically and, in fact, even use it predictively (Schremm et al., 2016). We added to this by analysing whether linguistic processing would emerge in a group of novice to intermediate learners.

Although the premises seemed supportive, we found no trace of rapid grammatical or predictive processing of Swedish grammatical tones for novice to intermediate learners. These results were unexpected. Our learners had studied Swedish for an average of 10 weeks (range: 0-24) and had lived in Sweden for 23 weeks (range: 7-78), (cf. Appendix A). We therefore predicted to find an N400 for tone mismatches, analogous to the previous findings of grammar acquisition in classroom settings (McLaughlin et al., 2010). The fact that no such effects were elicited might be due to the perceptually difficult tonal component. Previous studies have indicated general difficulties in the perception of L2 tone, especially for learners from a non-tonal L1 (e.g., Shen & Froud, 2019; Yang & Chan, 2010; Yu et al., 2019). Ambiguities in the perception and classification of L2 tone might thus hinder the detection and acquisition of tone-suffix rules.

Interestingly, we observed a sensitisation to the acoustic properties of the tones with increasing proficiency. This response is presumably related to preparatory processes of familiarisation and might be a required precursory step for the subsequent acquisition of the tones. Such familiarisation processes may even be mandatory for the acquisition of unfamiliar L2 input overall, particularly enhanced in the context of perceptually difficult features. Conclusively, the natural language acquisition results suggest that not all parts of grammar can be acquired with just minimal instruction, even if they are very frequent in the input language. Some features might rely on precursory steps.

5.2.2 Artificial language learning

While paper I examined the speed of the acquisition of L2 grammatical tone in a natural SLA context, papers II and III studied this in more controlled, artificial acquisition settings. We demonstrated that, behaviourally, learners showed most noticeable improvements of response accuracy and response times within 40 minutes of acquisition. For mismatches in the word-picture pairs, we found that learners produced an N400 within 60 minutes. During learning, on the other hand, grammatical processing and evidence of memory trace formation for the tonal words emerged within 20 minutes.

Previous studies of artificial L2 learning found very rapidly emerging P600 effects, often within a single acquisition session but at the latest after a 24-hour consolidation period (e.g., Batterink & Neville, 2013; de Diego-Balaguer et al., 2007; Havas et al., 2017). In contrast, our learners produced only the precursory N400 in response to violations in the scope of two 120-minute long acquisition sessions spaced 24 hours apart. A shift from the general, semantic mismatch-related N400 to the grammatical P600 was not observed, neither for tone nor vowel mismatches. The emergence of an N400 rather than a P600, which was found in comparable studies (e.g., Havas et al., 2017), was surprising as learners had arguably extracted the rules of the auditory words, see results of paper III. Evidently, then, the rules were not extended to or inaccessible for mismatches in picture referents. This may potentially be the case because learners had not surpassed an initial acquisitional stage where mismatch detection was more important than mismatch classification, as suggested in section 2.3.2.5.

Yet, comparing the N400 findings to the acquisition of grammatical tone in a natural language above, the observed processing in the artificial settings is relatively more advanced. This allows for the suggestion that natural second language learners might benefit from focused training of tones in grammatical tone languages to more quickly achieve native-like processing and tone awareness. A suggestion that has since been supported for Swedish grammatical tones whose rule-based processing elicited a LAN component in beginner to intermediate second language learners after tone-suffix training and improved tone production (Hed et al., 2019; Schremm et al., 2017). Focused training might, theoretically, improve the processing and acquisition of any difficult grammatical L2 feature.

The results for learning, in contrast to the mismatch results and to findings in previous papers on L2 grammar processing, suggest that the progression of grammar processing might be faster than is traditionally observed. Outside of the scope of violations, instead comparing newly acquired grammatical words to repeated non-words, grammatical ERP responses, including the preconscious ELAN could be evoked within just 20 minutes of acquisition. This is, to my knowledge, the earliest emergence of the ELAN and even the LAN that has been reported in the context of L2 words. Previous studies which looked at learning effects in artificial learning contexts (late vs early learning phases), reported N400 effects (de Diego-Balaguer et al., 2007; François et al., 2017) as well as earlier components (e.g., P2). Interestingly, the N400 learning effects in previous studies typically had an anterior distribution. It is theoretically possible that the components' anterior distribution was shaped by grammar processing in a similar manner to (left) anterior negativities.

However, while the learning results showed astonishingly fast grammatical processing of L2 words, which is an important finding in itself, they have no direct point of comparison in the literature. We studied processing outside of violation contexts, importantly, with strong control conditions to correct for effects of familiarisation or other factors that might potentially change over the course of the experimental training (e.g., motivation, tiredness, etc.). It is, therefore, difficult to discuss the relative speed of tone processing on the basis of the learning results and, consequentially, most reasonable to confine that discussion to the mismatch results.

In summary, part of current thesis used the case of grammatical tone to contribute to the discussion of how rapid second language grammar can be processed. To this effect, it was illustrated that the processing of mismatches in grammatical tone is delayed compared to violations of other grammatical features (e.g., de Diego-Balaguer et al., 2007; Havas et al., 2017; McLaughlin et al., 2010) in both natural and artificial learning settings. Crucially, we saw no transition from N400 to P600 for tone-suffix mismatches in the beginner to intermediate learners. In fact, we observed no mismatch responses in this context of natural tone. In the artificial paradigm, we did observe an N400 response to the tone mismatches but no P600 which is typically found very early for artificial learning. This is likely related to the perceptual difficulty of the words' tonal component which might slow the acquisition of all aspects of the L2 words even in controlled artificial settings. Support for the assumed impact of tone perception difficulty can be found in the relatively low accuracy of tone mismatch detection compared to other conditions even in learners who were familiar with tone from their native language. Thus, the acquisition and processing of new grammatical features can be rapid, but the degree of rapidity depends among other things on perception-related factors.

5.2.3 Limitations

With respect to the findings for natural grammatical tone in Swedish, it should be mentioned that Swedish tones and their associations with grammar are not explicitly taught in second (or even native) language classrooms. Yet, they are abundant in the input, as they are realised on every spoken word and implicit acquisition should be possible. Whether the types of grammatical features studied in McLaughlin et al. (2010) were explicitly taught is not clear. In theory, it is possible that the presence of explicit instruction might affect the speed of L2 processing. If this is the case, the observed delay in the L2 processing of Swedish tones might potentially be due to lack of instruction rather than perception difficulties. Alternatively, the results might even emerge in combination of both perceptual difficulties and lack of instruction.

It is further important to mention that we could not compare the canonical processing of tone in natural and artificial settings. The natural tones, which were contrasted with each other, differ in predictive strength but not with respect to their grammatical properties. Assuming that grammatical components are also elicited in non-violation contexts, both tones would equally strongly elicited grammatical processing on the basis of the word formation process that they take part in. Since the strength of their grammatical involvement does not differ, no differential grammatical response should emerge in comparison of Swedish tones. Instead, the natural tones have repeatedly been seen to elicit a component related to differentially strong pre-activation of upcoming language material in native speakers (Roll, 2015; Roll et al., 2015, 2017; Söderström, Horne, Frid & Roll, 2016). Since predictive processing is suggested to be very rare in L2 learners (e.g., Guillelmon & Grosjean, 2001), it was no great surprise that we did not find a PrAN component for tone in the beginner to intermediate learners. While a direct comparison of canonical processing in natural and artificial stimuli would be worthwhile, the present paradigms did not allow for that possibility.

5.3 Processing of violated and non-violated grammar

A final topic that was touched upon in the present thesis is related to potential differences between grammar processing in violation and non-violation contexts. As mentioned in the previous sections, we observed clear differences between the two types of contexts for artificial second language acquisition: Grammar violations, here expressed in mismatches between the grammatical content of words and a picture referent, elicited an N400 response. Finding an N400 effect rather than a P600 potentially indicated that learners were at a stage where they could detect the observed error but not yet identify it as grammatical (McLaughlin et al., 2010; Osterhout et al., 2008; Sorace, 1985). The non-violated processing of the grammatical words themselves, in contrast, elicited not only a P600 but also a LAN and even an ELAN in transfer-friendly settings. These findings illustrate two important things: Firstly, based on the comparison of processing in violation and non-violation contexts, grammar processing in second language learners might already take place before learners can correctly classify or even detect L2 errors. Following this argumentation, grammar processing might often go unnoticed in L2 processing studies. Secondly, the data confirms the common assumption that the grammatical components, observed traditionally in violation contexts, are in fact also present in and thus represent canonical grammar processing.

5.3.1 Differences between canonical and non-canonical processing

In the comparison of the results of papers II and III, it becomes evident that there might potentially be fundamental differences in how L2 learners process canonical grammar and grammar mismatches. The observed N400 for mismatch processing, on the one hand, is suggestive of a relatively early stage of L2 acquisition (cf. Sorace, 1985; McLaughlin et al., 2010). The data related to canonical processing, on the other hand, suggests full grammatical access within 20 minutes. Particularly the rapid emergence of LAN and ELAN responses are traditionally considered impossible at such an early stage. Even in artificial settings, they have at the very earliest been observed after several days of acquisition (Friederici et al., 2002; Morgan-Short et al., 2012). I suggest that the fully grammatical processing in canonical contexts, unexpected given the non-grammar-specific processing of mismatches, is related to the fact that we analysed non-violated stimuli. Grammar violations are more effortful to process than non-violated input (Hahne & Friederici, 2001; Perani & Abutalebi, 2005). This assumption is supported by the fact that we traditionally observe amplified ERP components, signalling increased neural activity, in response to violated grammar. Since L2 processing generally requires greater neural activity than L1 processing (Hahne & Friederici, 2001; Hasegawa et al., 2002), the available computational resources for error processing are presumably limited. In contexts then, where there are no mismatches, more resources are likely available leading to the elicitation of grammatical processing. This interpretation of the findings of the present dissertation encourages a careful consideration of the role that violations might play in the elicitation of grammatical processing. While violation paradigms have a strong tradition of being used to evoke language-related ERP components, they might not necessarily constitute the best method for studying grammar processing in strenuous conditions. This is especially true for second language learners. Interestingly, a number of recent studies on rapid L2 morphosyntax acquisition have reported learning effects in addition to violation effects (e.g., Cunillera et al., 2009; de Diego-Balaguer et al., 2007; François et al., 2017). These studies consistently reported frontal negativities which might be related to the frontal negativities that we observed for learning effects. Given the assumed relative difficulty in the detection and classification of L2 errors, violation-free processing might show considerably earlier onsets of L2 grammar processing than what is thus far accepted.

An alternative explanation of the observed differences between papers II and III can be found in methodological choices. In the word-picture association paradigm, violations were realised in picture referents rather than words. Although picture processing has previously been seen to evoke a late LAN response in grammatical contexts (Wicha et al., 2003), the visual form can have affected the learners' processing of the mismatches. Learners could have responded to semantic information in the pictures, e.g., LONG HAIR or FEMALE, rather than the highly abstract grammatical feature FEMININE. Although the grammatical processing of the words themselves illustrated that our learners were able to correctly abstract the grammatical categories and associate them with the pseudowords, they may still have processed the same features in a more semantic manner in the context of the pictures. The use of picture referents in the mismatch paradigm might be the cause of or a contribution to the non- or pre-grammatical processing that was observed.

5.3.2 ERP components for canonical language

The comparison of paper II and paper III importantly also gives strong evidence for the claim that grammatical processing also occurs outside of violation contexts. This finding is particularly important in the field of language-related ERP research as it validates the underlying assumption that canonical grammar processing draws on the same resources as violation processing and that violation paradigms can consequentially illustrate general characteristics of language processing. The present results are amongst few to elicit full-fledged grammatical processing in canonical language. The P600 has often previously been elicited in fully grammatical sentences where local predictions did not hold. This was even the case, for instance, in the original P600 paper by Osterhout and Holcomb (1992) where the P600 was elicited for prepositions that motivated a reinterpretation of the sentence initial noun phrase as object rather than subject (i.e., garden path sentences). Relatively recently, a LAN response has also been observed in non-violated language. To this end, Krott and Lebib (2013) reported a LAN that was differently modulated by regular and irregular past participles. With respect to the ELAN, Steinhauer & Drury (2012) probably rightfully claim that none of the traditional ELAN elicitation paradigms are outright violations. All sentences can theoretically be continued grammatically at suggested word category violation points. However, both the P600 and the ELAN paradigms likely introduce local violations due to strong sentential predictions. As the inflected nouns used in the acquisition paradigm are not embedded in sentential contexts, such effects do not apply here. Having closely controlled the frequency and phonological structure of the newly acquired words and the pseudowords of the control condition, we were able to construct conditions where grammatical content was the most compelling difference, resulting in the elicitation of a full range of grammar components. This brings forward conclusive evidence of grammatical processing in non-violation contexts. It is important to stress, however, that the findings we observed here were for L2 learners. In this context, Hahne and Friederici (2001) argue that non-violated language for L2 learners might have the same processing load as violated language in native speakers. It is thus possible that the results are L2-specific and do not replicate in L1 contexts.

6 Conclusions

The present thesis investigated the role of phonological transfer on the initial stages of grammatical tone acquisition and processing in second language learners. Based on evidence from both natural and artificial second language settings, it illustrated that phonological similarity between native and target language strongly impacts the processing of L2 words in early stages of second language acquisition. The earlier the processing, the more strongly it depends on phonological L1-L2 similarity. Late processing and acquisition (as measured by behavioural responses), on the other hand, do not necessarily appear to be affected by transfer. Thus, while transfer crucially affects some parts of processing in the initial stages of second language acquisition, it may not strictly be necessary for successful acquisition. The lack of positive transfer can be compensated for by late, more conscious processing to such a degree that behavioural results can be indistinguishable. Further, comparing the acquisition of grammatical tone to that of comparable non-tonal grammatical features, the dissertation found a general acquisition and processing delay for L2 tone. This suggests that the speed of L2 acquisition seems to depend crucially on the difficulty (here likely perceptual) of the L2 input which is not necessarily overcome by transfer-based facilitation of preconscious processing. However, virtually instantaneous grammatical processing emerged for canonical processing of artificial grammatical words, suggesting that grammar processing in L2 learners may occur much earlier than is traditionally thought possible and that the effectiveness of violation paradigms in the discussion of L2 grammar processing might have to be carefully evaluated.

These findings have different implications for and contribute differently to the three communities for whom this dissertation is most relevant: ERP researchers, SLA researchers, and general linguists:

Importantly for ERP research, the present thesis produced evidence of grammatical ERP components in canonical contexts, providing strong support for the generally accepted but rarely studied assumption that processing of violations and processing of non-violated language involve the same neural processes. The findings were similar to but expanded on previous findings (e.g., Krott & Lebib, 2013) by producing not only a LAN in canonical contexts but also an ELAN in clearly non-violated contexts. This has extended the assumption that the ERP components for language, established on the basis of violation processing, do indeed reflect general neurophysiological processes which are strongly involved in language comprehension.

Directly comparing violation and non-violation contexts, the results in the present dissertation additionally have potential implications for the use of violation paradigms in language-based ERP studies. They suggest that violation paradigms, while reflective of core linguistic processes, might not be ideal for the study of grammar processing in resource-heavy conditions. This was found here for second language learners but might more generally extend to other, resource-heavy processing contexts.

The thesis further has implications for the interpretation of several ERP effects, in particular ELAN and LAN. Considering the ELAN, its elicitation in the context of single, inflected words urges an interpretation of the effect away from 'phrase structure' or 'word category' (Friederici et al., 1993; Neville et al., 1991) and towards 'grammaticality assessment' (cf. Pulvermüller & Assadollahi, 2007). Thus, the single word presentations indicate that sentential or phrasal contexts do not govern the ELAN. Instead, it is likely related to a preconscious activation and assessment of the words' rule-based, grammatical content. This suggested type of processing can easily explain differences between words with and without grammatical content, on the one hand, and words with and without grammatical violations, on the other: Preconscious processing of grammatical content elicits an ELAN response that is further increased as an indication of processing load in the presence of violations. Considering the LAN, the hitherto most frequent interpretations of its elicitation conditions are that it is strongly modulated by prediction (e.g., Wicha et al., 2004) or indicative of agreement processing (e.g., Molinaro et al., 2014). While both may be true, the present results suggest that congruity with context-based grammatical predictions or agreement cannot be the only factors differentially affecting the LAN response. The conditions which evoked LAN amplitude differences in the present dissertation had no differential predictive weight. Instead, they differed purely with respect to whether or not they contained grammatical affixes. A tentative interpretation of the LAN, encompassing previous and current results, might therefore be that it signals rule-based word-level integration of grammar and lexical semantic features which is facilitated for words that agree with and can be predicted from the context.

The dissertation also makes several contributions to the field of SLA. Firstly, and maybe most importantly, it informs about the intricacies of transfer in initial acquisition. Somewhat at odds with theories that claim that initial L2 acquisition is not mediated through the L1 (e.g., Pienemann, 1998), I have reason to argue that transfer, at least at the level of phonology, does in fact affect learners during initial acquisition. Yet, transfer appears to be most important at ultra-rapid processing levels and its impact decreases the later the processing. In consequence, transfer may not influence the latest, most consciously analysed component nor may the transfer-based facilitation of early processing manifest as behavioural acquisition differences. It is possible that the results in this study are influenced by ceiling effects and that late processing and behaviour might be differentially affected in more strenuous acquisition settings. In sum, the findings in the thesis allow for the conclusion that, while transfer effects are not always observable in the learners' behavioural responses, L1-L2 phonetic and phonological similarity presumably always plays an important role in how learners initially process L2 input. It is likely that this conclusion extends beyond the realm of phonology (for similar findings for syntactic transfer, see Andersson et al., 2019), and it implies that L1 and L2 processing are mediated through the same neural networks.

Further relevant for SLA is the invitation to be critical about (violation-based) grammaticality judgement data in learner populations. While they undoubtedly demonstrate interesting and relevant effects, it is possible that they are unable to adequately capture the true progression of acquisition. Learners might have internal concepts of L2 grammar features before they are able to notice mismatches. Not correctly responding to a mismatch does therefore not necessarily imply that the underlying structure is not yet (at least in part) acquired and used by the learner.

The findings in the thesis, finally, also have interesting implications for general linguistics, specifically for the study of phonetics and phonology. Results for the word recognition component suggested that only falling pitch contours of L2 word-level tone led to the formation of rapid memory traces. These neurophysiological findings provide strong experimental support for Bruce's characterisation of Swedish tones as falls with different timings with respect to the stressed vowel (Bruce, 1977, 1987, 2005). Presupposing that the amplitude differences observed in the word recognition component are related to the fact that novel words can make differential use of pre-existing neural networks, it is possible to argue that there are facilitative pre-existing networks only for tones with falling pitch. Such networks would be formed by the functional use of falling word-level tones in Swedish, which validates the phonetic description of falling pitch movements as the minimally distinctive and thus perceptually most important characteristic of tone in Swedish.

7 Outstanding issues and future directions

As mentioned in the discussion, there are some aspects of the thesis that call for further support to strengthen the postulated interpretations. One such thing became apparent in the mismatch processing results. Clear differential effects of phonological transfer were observed for mismatch processing, suggesting that the tones influenced the detection of mismatches in this context. Yet, it cannot with certainty be decided what generated the elicitation of the observed N400 component. This interpretation difficulty arises because mismatches were embedded in the pictorial referent of the word-picture pair. It is possible that grammatical processing was simply not possible in the context of the pictures and that this evoked the N400 response, as is the case for semantically incongruent pictures (Nigam et al., 1992). Alternatively, the N400 response shows a precursory stage of grammar processing often seen in L2 contexts (cf. McLaughlin et al., 2010). There are factors that point towards the latter interpretation, such as the fact that LAN-like responses have previously been observed in the context of picture processing (Wicha et al., 2003) or the fact that the pictures were able to install grammatical processing in the preceding words. Future studies are required to substantiate the interpretation. If one wanted to use a paradigm similar to the one in this dissertation, it would be possible to use word-word associations rather than word-picture associations. This way, mismatches would be embedded in the associated words rather than pictures, creating agreement-like conditions. However, this would increase the influence of the native language, if native words were used in the associations. Such influences might be best to avoid. Alternatively, one could include small test blocks throughout the acquisition paradigm, which would, however, possibly have an impact on learning and add explicit focus on the mismatches.

The experimental implementation of mismatches in picture referents has an additional disadvantage: it undermines the comparison of violated and non-violated language. While there were clear differences in the learning effects and the mismatch effects, these could potentially be attributed in full to the fact that pictures were used to elicit mismatch processing. As argued above, based on the observed response patterns and with respect to previous results, I view it as unlikely that the processing of the pictures was altogether unaffected by the grammatical categories. In the theoretical scenario that this was the case, there is still a noticeable clash between the unprecedentedly rapid emergence of early, preconscious processing in canonical learning contexts, especially considering that we looked at a group of second language learners, and the low behavioural accuracy for mismatch detection (below 70%). Components like the LAN and the ELAN typically do not appear until much higher accuracy in artificial learning studies with violation paradigms (Friederici et al., 2002: 95% threshold; Morgan-Short et al., 2010: >90% average, minimum 80%). This strengthens the suggestion that there is a fundamental difference between mismatch processing and canonical processing in learners. The case of canonical processing in the present study thus suggests that learners might be capable of grammar processing in L2 contexts much earlier than they are believed to. Yet, irrespective of how the results are to be interpreted, the current findings for canonical processing of L2 grammar suggest the usefulness of a systematic investigation of canonical processing also in natural second language learners, comparing, for instance, inflected pseudowords to pseudo-inflected pseudowords. Such studies have the potential to provide important insights into learner processing and possible even language processing overall, if extended to L1 contexts.

An alternative explanation of the observed differences between violation and non-violation processing (if not a picture-processing artefact) might be that L2 processing fundamentally differs from the L1 with respect to processing load (Hahne & Friederici, 2001). Thus, L2 words in initial learners might have an increased processing load similar to violation processing in the L1. If this is the case, canonical L1 processing could be considered the least resource-heavy ‘default’ condition which would be affected equally strongly with respect to increased processing load in violation contexts and for L2 processing. While possible to study the validity of this possibility with the help of EEG and differential amplitude differences in the LAN component, it could also be targeted, for instance, with functional magnetic resonance imaging (fMRI). Canonical language, violations as well as L2 processing are argued to engage the same core language networks with different intensity (Mollica et al., 2020; Perani & Abutalebi, 2005). As a result, relative activation of those core networks in early L2 learners and native speakers with respect to both grammatical and ungrammatical language, could inform about potential quantitative differences between L1 and L2 processing, on the one hand, and canonical and violation processing, on the other.

Future research should also more closely target the role of falling tones in Swedish. While the present results were a first attestation of the phonological reality of the Swedish tones, they were, admittedly, unexpected and the elicitation conditions were not ideal. With more carefully configured experimental settings, testing for instance the influence of using more and less similar pitch slopes and more deliberately comparing falls to rises, it might be possible to tap more into the functional importance of falling pitch at the word level in Swedish.

References

Andersson, A., Sayehli, S., & Gullberg, M. (2019). Language background affects online word order processing in a second language but not offline. *Bilingualism: Language and Cognition, 22*(4), 802-825. https://doi.org/10.1017/S1366728918000573

Bakker, I., MacGregor, L. J., Pulvermüller, F., & Shtyrov, 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 H. van der Hülst, & N. Smith (Eds.), *Autosegmental studies on pitch accent* (pp. 11-50). Foris Publications.

Batterink, L., & Neville, H. (2013). Implicit and explicit second language training recruit common neural mechanisms for syntactic processing. *Journal of Cognitive Neuroscience, 25*(6), 936-951. https://doi.org/10.1162/jocn_a_00354.

Bauer, L. (2003). *Introducing Linguistic Morphology* (2nd ed.). Edinburg University Press.

Blomberg, F., Roll, M., Frid, J., Lindgren, M., & Horne, M. (2020). The role of affective meaning, semantic associates, and orthographic neighbours in modulating the N400 in single words. *The Mental Lexicon, 15*(2), 159-186.

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

Braun, B., & Johnson, E. K. (2011). Question or tone 2? How language experience and linguistic function guide pitch processing. *Journal of Phonetics, 39*, 585-594. https://doi.org/10.1016/j.wocn.2011.06.002

Brown-Schmidt, S., & Canseco-Gonzalez, E. (2004). Who do you love, your mother or your horse? An event-related potential analysis of tone processing in Mandarin Chinese. *Journal of Psycholinguistic Research, 33*(2), 103-135. https://doi.org/10.1023/B:JOPR.0000017223.98667.10

Bruce, G. (1977). *Swedish word accents in sentence perspective.* Gleerup.

Bruce, G. (1987). How floating is focal accent? In K. Gregersen, & H. Basbøll (Eds.), *Nordic prosody IV: papers from a symposium* (pp. 41-49). Odense University Press.

Bruce, G. (2005). Intonational prominence in varieties of Swedish revisited. In S.-A. Jun (Ed.), *Prosodic typology: The phonology of intonation and phrasing.* https://doi.org/10.1093/acprof:oso/9780199249633.003.0015

Bruce, G., & Gårding, E. (1978). *A prosodic typology for Swedish dialects.* (E. Gårding, G. Bruce, & R. Bannert, Eds.). Department of Linguistics, Lund University.

Burnham, D., Kasisopa, B., Reid, A., Luksaneeyanawin, S., Lacerda, F., Attina, V., Rattanasone, N. X., Schwarz, I.-C., & Webster, D. (2015). Universality and language-specific experience in the perception of lexical tone and pitch. *Applied Psycholinguistics, 36*, 1459-1491. https://doi.org/10.1017/S0142716414000496

Bybee, J., & Newman, J. E. (1995). Are stem changes as natural as affixes? *Linguistics, 33*(4), 633-654. https://doi.org/10.1515/ling.1995.33.4.633

Carrasco-Ortíz, H., Herrera, A. V., Jackson-Maldonado, D., Avecilla Ramírez, G. N., Silva Pereyra, J., & Wicha, N. Y. (2017). The role of language similarity in processing second language morphosyntax: Evidence from ERPs. *International Journal of Psychophysiology, 117*, 91-110. https://doi.org/10.1016/j.ijpsycho.2017.04.008

Chan, R. K., & Leung, J. H. (2020). Why are lexical tones difficult to learn?: Insights from the incidental learning of tone-segment connections. *Studies in second language acquisition, 42*(1), 33-59. https://doi.org/10.1017/S0272263119000482

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, B., Ma, T., Liang, L., & Liu, H. (2017). Rapid L2 word learning through high constraint sentence context: An event-related potential study. *frontiers in Psychology, 8*(2285). https://doi.org/10.3389/fpsyg.2017.02285

Chomsky, N. (1965). *Aspects of the theory of syntax.* M.I.T. Press.

Collentine, J., & Freed, B. F. (2005). *Learning context and its effects on second language acquisition* (Vol. 26). Cambridge University Press. Retrieved from https://www.jstor.org/stable/44486767

Crysmann, B. (2015). Representing morphological tone in a computational grammar of Hausa. *Journal of Language Modelling, 3*(2), 463-512. https://doi.org/10.15398/jlm.v3i2.126.

Cunillera, T., Càmara, E., Toro, J. M., Marco-Pallares, J., Sebasián-Galles, N., Ortiz, H., Pujol, J, & Rodríguez-Fornells, A. (2009). Time course and functional neuroanatomy of speech segmentation in adults. *NeuroImage*, *48*(3), 541-553. https://doi.org//10.1016/j.neuroimage.2009.06.069

Davis, M. H., Di Betta, A. M., Macdonald, M. J., & Gaskell, M. G. (2009). Learning and consolidation of novel spoken words. *Journal of Cognitive Neuroscience, 21*(4), 803-820. https://doi.org/10.1162/jocn.2009.21059.

de Diego-Balaguer, R., Toro, J. M., Rodríguez-Fornells, A., & Bachoud-Lévi, A.-C. (2007). Different neurophysiological mechanisms underlying word and rule extraction from speech. *PLoS ONE*, *11*. Article e1175. https://doi.org/10.1371/journal.pone.0001175

Dien, J. (1998). Issues in the application of the average reference: Review, critique, and recommendations. *Behavior Research Methods, Instruments, & Computers, 30*(1), 34-43. https://doi.org/10.3758/BF03209414

Dittinger, E., Scherer, J., Jäncke, L., Besson, M., & Elmer, S. (2019). Testing the influence of musical expertise on novel word learning across the lifespan using a cross-sectional approach in children, younger adults and older adults. *Brain and Language, 198*(104678). https://doi.org/10.1016/j.bandl.2019.104678

Eberhard, D. M., Simons, G. F., & Fennig, C. D. (Eds.). (2020). *Ethnologue: Languages of the world* (23rd ed.). SIL International. Online version: http://www.ethonologue.com.

Enger, H.-O. (2005). The Nordic languages in the 19th century II: Morphology. In O. Bandle, K. Braunmüller, E.-H. Jahr, H.-P. Naumann, A. Karker, & U. Teleman (Eds.), *The Nordic languages: An international handbook of the history of the North Germanic languages* (pp. 1437-1442). Walter de Gruyter.

Ettlinger, M., Morgan-Short, K., Faretta-Stutenberg, M., & Wong, P. C. (2016). The relationship between artificial and second language learning. *Cognitive Science, 40*, 822-847. https://doi.org/10.1111/cogs.12257

Féry, C. (1994). *Umlaut and inflection in German.* Department of Linguistics, University of Potsdam.

Foucart, A., & Frenck-Mestre, C. (2012). Can late L2 learners acquire new grammatical features? Evidence from ERPs and eye-tracking. *Journal of Memory and Language, 66*, 226-248. https://doi.org/10.1016/j.jml.2011.07.007

François, C., Cunillera, T., Garcia, E., Laine, M., & Rodriguez-Fornells, A. (2017). Neurophysiological evidence for the interplay of speech segmentation and word-referent mapping during novel word learning. *Neuropsychologia*, 56-67. https://doi.org/10.1016/j.neuropsychologia.2016.10.006

Friederici, A. D., Pfeifer, E., & Hahne, A. (1993). Event-related brain potentials duuring naturual speech processing: Effects of semantic, morphological and syntactic violations. *Cognitive Brain Research, 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

Gandour, J. (1983). Tone perception in Far Eastern languages. *Journal of Phonetics, 11*, 149-175. https://doi.org/10.1016/S0095-4470(19)30813-7

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

Gårding, E., & Lindblad. (1973). Constancy and variation in Swedish word accent patterns. *Working Papers* (Phonetics Laboratory, Lund University)*, 7*, 36-110.

Genesee, F. (1994). Integrating language and content: Lessons from immersion.

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

Gillon Dowens, M., Guo, T., Guo, J., Barber, H., & Carreiras, M. (2011). Gender and number processing in Chinese learners of Spanish - Evidence from event related potentials. *Neuropsychologia, 49*, 1651-1659. https://doi.org/10.1016/j.neuropsychologia.2011.02.034

Gillon Dowens, M., Vergara, M., Barber, H. A., & Carreiras, M. (2010). Morphosyntactic processing in late second-language learners. *Journal of Cognitive Neuroscience, 22*(8), 1870-1887. https://doi.org/10.1162/jocn.2009.21304

Gosselke Berthelsen, S., Horne, M., Brännström, K. J., Shtyrov, Y., & Roll, M. (2018). Neural processing of morphosyntactic tonal cues in second-language learners. *Journal of Neurolinguistics, 45*, 60-78. https://doi.org/10.1016/j.jneuroling.2017.09.001

Guillelmon, D., & Grosjean, F. (2001). The gender marking effect in spoken word recognition: The case of bilinguals. *Memory & Cognition, 29*(3), 503-511. https://doi.org/10.3758/bf03196401

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. (2012). What word-level knowledge can adult learners acquire after minimal exposure to a new language? *International Review of Applied Linguistics in Language Teaching, 50*, 239-276. https://doi.org/10.1515/iral-2012-0010

Gussenhoven, C. (2004). *The phonology of tone and intonation.* Cambridge University Press.

Hahne, A. (2001). What's different in second-language processing? Evidence from event-related brain potentials. *Journal of Psycholinguistic Research, 30*(3), 251-266. https://doi.org/10.1023/A:1010490917575

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

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. (2018). Second language perception of Mandarin vowels and tones. *Language and Speech, 61*(1), 135-152. https://doi.org/10.1177/0023830917717759

Hasegawa, M., Carpenter, P. A., & Just, M. A. (2002). An fMRI study of bilingual sentence comprehension and workload. *NeuroImage, 15*, 647-660. https://doi.org/10.1006/nimg.2001.1001

Hasting, A. S., & Kotz, S. A. (2008). Speeding up syntax: on the relative timing and automaticity of local phrase structure and morphosyntactic proessing as reflected in event-related brain potentials. *Journal of Cognitive Neuroscience, 20*(7), 1207-1219. https://doi.org/10.1162/jocn.2008.20083

Havas, V., Laine, M., & Rodriguez-Fornells, A. (2017). Brain signatures of early lexical and morphological learning of a new language. *Neuropsychologia, 101*, 47-56. https://doi.org/10.1016/j.neuropsychologia.2017.04.005

Havas, V., Taylor, J., Vaquero, L., de Diego-Balaguer, R., Rodríguez-Fornells, A., & Davis, M. H. (2018). Semantic and phonological schema influence spoken word learning and overnight consolidation. *Quarterly Journal of Experimental Psychology, 71*(6), 1469-1481. https://doi.org/10.1080/17470218.2017.1329325

Hed, A., Schremm, A., Horne, M., & Roll, M. (2019). Neural correlates of second language acquisition of tone-grammar associations. *The Mental Lexicon, 14*(1), 98-123. https://doi.org/10.1075/ml.17018.hed

Herrmann, B., Maess, B., & Friederici, A. D. (2011). Violation of syntax and prosody - Disentangling their contributions to the early left anterior negativity. *Neuroscience Letters, 490*, 116-120. https://doi.org/10.1016/j.neulet.2010.12.039

Höder, S. (2016). Phonological elements and diasystematic construction grammar. In M. Hilpert, & J.-O. Östman (Eds.), *Constructions across grammars.* John Benjamins.

Huang, T., & Johnson, K. (2010). Language specificity in speech perception: Perception of Mandarin tones by native and nonnative listeners. *Phonetica, 67*, 243-267. https://doi.org/10.1159/000327392

Hyman, L. M. (2016). Lexical vs. grammatical tone: Sorting out the differences. *Proceedings of Tonal Aspects of Languages 2016*, 6-11. https://doi.org/10.21437/TAL.2016-2

Jung, T.-P., Makeig, S., Humphries, C., Lee, T.-W., McKeown, M. J., Iragui, V., & Sejnowski, T. J. (2000). Removing electroencephalographic artifacts by blind source separation. *Psychophysiology, 37*, 163-178. https://doi.org/10.1111/1469-8986.3720163

Kim, A., & Osterhout, L. (2005). The independence of combinatory semantic processing: Evidence from event-related potentials. *Journal of Memory and Language, 52*(2), 205-225. https://doi.org/10.1016/j.jml.2004.10.002

Kimppa, L., Kujala, T., Leminen, A., Vainio, M., & Shytrov, 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., Shtyrov, Y., Hut, S. C., Hedlund, L., Leminen, M., & Leminen, A. (2019). Acquisition of L2 morphology by adult language learners. *Cortex, 116*, 74-90. https://doi.org/10.1016/j.cortex.2019.01.012

Kluender, R., & Kutas, M. (1993). Bridging the gap: evidence from ERPs on the processing of unbounded dependencies. *Journal of Cognitive Neuroscience, 5*(2), 196-214. https://doi.org/10.1162/jocn.1993.5.2.196

Kochančikaitė, R., Shtyrov, Y., & Roll, M. (2020). Lexical and decompositional processing in parallel? Evidence from Swedish, a pitch-accent language. Manuscript in preparation.

Koelsch, S., Gunter, T., Friederici, A. D., & Schröger, E. (2000). Brain indices of music processing: "Nonmusicians" are musical. *Journal of Cognitive Neuroscience, 12*(3), 520-541. https://doi.org/10.1162/089892900562183

Koelsch, S., Schmidt, B.-H., & Kansok, J. (2002). Effects of musical expertise on the early right anterior negativity: An event-related brain potential study. *Psychophysiology, 39*, 657-663. https://doi.org/10.1017.S0048577202010508

Kotz, S. A. (2009). A critical review of ERP and fMRI evidence on L2 syntactic processing. *Brain & Language, 109*, 68-74. https://doi.org/10.1016/j.bandl.2008.06.002

Krashen, S. D., Long, M. A., & Scarcella, R. C. (1979). Age, rate and eventual attainment in second language acquisition. *TESOL Quarterly, 13*(4), 573-582. https://doi.org/10.2307/3586451

Krott, A., & Lebib, R. (2013). Electrophysiological evidence for a neural substrate of morphological rule application in correct wordforms. *Brain Research, 1496*, 70-83. https://doi.org/10.1016/j.brainres.2012.12.012

Kutas, M., & Federmeier, K. D. (2011). Thirty years and counting: Finding meaning in the N400 component of the event related brain potential (ERP). *Annual Review of Psychology, 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. https://doi.org/10.1017/S0008413100025196

Lee, Y.-S., Vakoch, D. A., & Wurm, L. H. (1996). Tone Perception in Cantonese and Mandarin: A cross-linguistic comparison. *Journal of Psycholinguistic Research, 25*(5), 527-542. https://doi.org/10.1007/BF01758181

Lehmann, C. (1995). *Thoughts on grammaticalization.* Lincom Europa.

Li, X., Yang, Y., & Hagoort, P. (2008). Pitch accent and lexical tone processing in Chinese discourse comprehension: An ERP study. *Brain Research, 1222*, 192-200. https://doi.org/10.1016/j.brainres.2008.05.031

Luck, S. J. (2005). *An introduction to the event-related potential technique.* MIT Press.

Luck, S. J. (2014). *An introduction to the event-related potential technique* (2nd ed.). MIT Press.

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. *Language Learning, 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

Malins, J. G., & Joanisse, M. F. (2012). Setting the tone: An ERP investigation of the influences of phonological similarity on spoken word recognition in Mandarin Chinese. *Neuropsychologia, 50*, 2032-2043. https://doi.org/10.1016/j.neuropsychologia.2012.05.002

Marantz, A. (1997). No escape from syntax: Don't try morphological analysis in the privacy of your own lexicon. *Unviesrity of Pennsylvania Working Papers in Linguistics, 4*(2), 200-225.

Marcus, G. F., Brinkmann, U., Clahsen, H., Wiese, R., & Pinker, S. (1995). German inflection: The exception that proves the rule. *Cognitive Psychology, 29*(3), 189-256. https://doi.org/10.1006/cogp.1995.1015

McLaughlin, J., Osterhout, L., & Kim, A. (2004). Neural correlates of second-language word learning: minimal instruction produces rapid change. *Nature Neuroscience, 7*(7), 703-704. https://doi.org/10.1038/nn1264

McLaughlin, J., Tanner, D., Pitkänen, I., Frenck-Mestre, C., Inoue, K., Valentine, G., & Osterhout, L. (2010). Brain potentials reveal discrete stages of L2 grammatical learning. *Language Learning, 60*(Suppl. 2), 123-150. https://doi.org/10.1111/j.1467-9922.2010.00604.x

Meisel, J. M. (2009). Second language acquisition in early childhood. *Zeitschrift für Sprachwissenschaft, 28*, 5-34. https://doi.org/10.1515/ZFSW.2009.002

Midgley, K. J., Holcomb, P. J., & Grainger, J. (2011). Effects of cognate status on word comprehension in second language learners: An ERP investigation. *Journal of Cognitive Neuroscience, 23*(7), 1634-1647. https://doi.org/10.1162/jocn.2010.21463

Molinaro, N., Barber, H. A., Caffarara, S., & Carreiras, M. (2014). On the left anterior negativity (LAN): The case of morphosyntactic agreement. *Cortex*, 1-4. https://doi.org/10.1016/j.cortex.2014.06.009

Mollica, F., Siegelman, M., Diachek, E., Piantadosi, S. T., Mineroff, Z., Futrell, R., Qian, P., & Fedorenko, E. (2020). Composition is the core driver of the language-selective network. *Neurobiology of Language, 1*(1), 104-134. https://doi.org/10.1162/nol_a_00005

Morgan-Short, K., Sanz, C., Steinhauer, K., & Ullman, M. T. (2010). Second language acqusition of gender agreement in explicit and implicit training conditions: An event-related potential study. *Language Learning, 60*(1), 154-193. https://doi.org/10.1111/j.1467-9922.2009.00554.x

Morgan-Short, K., Steinhauer, K., Sanz, C., & Ullman, M. T. (2012). Explicit and implicit second language training differentially affect the achievement of native-like brain activation patterns. *Journal of Cognitive Neuroscience, 24*(4), 993-947. https://doi.org/10.1162/jocn_a_00119

Mueller, J. L., Hahne, A., Fujii, Y., & Friederici, A. D. (2005). Native and nonnative speakers' processing of a miniature version of Japanese as revealed by ERPs. *Journal of Cognitive Neuroscience, 17*(8), 1229-1244. https://doi.org/10.1162/0898929055002463

Mueller, J. L., Hirotani, M., & Friederici, A. D. (2007). ERP evidence for different strategies in the processing of case markers in native speakers and non-native learners. *BMC Neuroscience, 8*(18). https://doi.org/10.1162/0898929055002436

Neville, H., Nicol, J. L., Barss, A., Forster, K. I., & Garret, M. F. (1991). Syntactically based sentence processing classes: Evidence from event-related potentials. *Journal of Cognitive Neuroscience, 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

Nigam, A., Hoffman, J. E., & Simons, R. F. (1992). N400 to semantically anomalous pictures and words. *Journal of Cognitive Neuroscience, 4*(1), 15-22. https://doi.org/10.1162/jocn.1992.4.1.15

Oomen, A. (1981). Gender and plurality in Rendille. In R. Hetzron, & R. G. Schuh (Eds.), *Afroasiatic Linguistics* (Vol. 8, pp. 35-75). Undenoa.

Osterhout, L., & Holcomb, P. J. (1992). Event-related brain potentials elicited by syntactic anomaly. *Journal of Memory and Language, 31*, 785-806. https://doi.org/10.1016/0749-596X(92)90039-Z

Osterhout, L., & Mobley, L. A. (1995). *Journal of Memory and Language, 34*(6), 739-773. https://doi.org/10.1006/jmla.1995.1033

Osterhout, L., Poliakov, A., Inoue, K., McLaughlin, J., Valentine, G., Pitkanen, I., Frenck-Mestre, C., & Hirschensohn, J. (2008). Second-language learning and changes in the brain. *Journal of Neurolinguistics, 21*(6), 509-521. https://doi.org/10.1016/j.jneuroling.2008.01.001

Paradis, M. (2004). A neurolinguistic theory of bilingualism. (K. de Bot, & T. Huebner, Eds.) John Benjamins.

Pelzl, E., Lau, E. F., Guo, T., & DeKeyser, R. (2019). Advanced second language learners' perception of lexical tone contrasts. *Studies in Second Language Acquisition, 41*, 59-86. https://doi.org/10.1017/S0272263117000444

Pelzl, E., Lau, E. F., Guo, T., & DeKeyser, R. (2020). Even in the best-case scenario L2 learners have persistent difficulty perceiving and utilizing tones in Mandarin: Findings from behavioural and event-related potentials experiments. *Studies in Second Language Acquisition*, 1-29. https://doi.org/10.1017/S027226312000039X

Perani, D., & Abutalebi, J. (2005). The neural basis of first and second language processing. *Current Opinion in Neurobiology*, 202-206. https://doi.org/10.1016/j.conb.2005.03.007

Pienemann, M. (1998). Language processing and second language development: Processability theory. In K. de Bot, & T. Huebner (Eds.), *Studies in bilingualism.* John Benjamins.

Pulvermüller, F., & Assadollahi, R. (2007). Grammar or serial order?: Discrete combinatorial brain mechanisms reflected by the syntactic mismatch negativity. *Journal of Cognitive Neuroscience, 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

Pulvermüller, F., Cappelle, B., & Shrytov, Y. (2013). Brain basis of meaning, words, constructions, and grammar. In T. Hoffmann, & G. Trousdale (Eds.), *The Oxford Handbook of Construction Grammar.* Oxford University Press. https://doi.org/10.1093/oxfordhb/9780195396683.013.0022

Rast, R. (2010). The role of linguistic input in the first hours of adult language learning. *Language Learning, 60*(Suppl. 2), 64-84.

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

Richards, J. C. (2015). The changing face of language learning: learning beyond the classroom. *RELC, 46*(1), 5-22. https://doi.org/10.1177/0033688214561621

Rodriguez-Fornells, A., Clahsen, H., Lleó, C., Zaake, W., & Münte, T. F. (2001). Event-related brain responses to morphological violations in Catalan. *Cognitive Brain Research, 11*(1), 47-58. https://doi.org/10.1016/S0926-6410(00)00063-X

Roll, M. (2015). A neurolinguistic study of South Swedish word accents: Electrical brain potentials in nouns and verbs. *Nordic Journal of Linguistics, 38*(2), 149-162. https://doi.org/10.1017/S0332586515000189

Roll, M., Gosselke, S., Lindgren, M., & Horne, M. (2013). Time-driven effects on processing grammatical agreement. *Frontiers in Psychology, 4*(1004), 1-8. https://doi.org/10.3389/fpsyg.2013.01004

Roll, M., Söderström, P., & Horne, M. (2013). Word-stem tones cue suffixes in the brain. *Brain Research, 1520*, 116-120. https://doi.org/10.1016/j.brainres.2013.05.013

Roll, M., Söderström, P., Frid, J., Mannfolk, P., & Horne, M. (2017). Forehearing words: Pre-activation of word endings at word onset. *Neuroscience Letters, 658*, 57-61. https://doi.org/10.1016/j.neulet.2017.08.030

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 & Language, 150*, 14-21. https://doi.org/10.1016/j.bandl.2015.07.009

Schremm, A., Hed, A., Horne, M., & Roll, M. (2017). Training predictive L2 processing with a digital game: Prototype promotes acquisition of anticipatory use of tone-suffix associations. *Computers & Education, 114*, 206-221. https://doi.org/10.1016/j.compedu.2017.07.006

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

Schremm, A., Söderström, P., Horne, M., & Roll, M. (2016). Implicit acquisition of tone-suffix connections in L2 learners of Swedish. *The Mental Lexicon, 11*(1), 55-75. https://doi.org/10.1075/ml.11.1.03sch

Selinker, L. (1988). *Papers in Interlanguage. Occasional Papers No. 44.* SEAMEO Regional Language Centre.

Shen, G., & Froud, K. (2019). Electrophysiological correlates of categorical perception of lexical tones by English learners of Mandarin Chinese: an ERP study. *Bilingualism: Language and Cognition, 22*(2), 253-265. https://doi.org/10.1017/S136672891800038X

Shih, C., & Lu, H.-Y. D. (2010). Prosody transfer and suppression: Stages of tone acquisition. *Speech Prosody 2010*, *968.*

Shtyrov, Y., Pulvermüller, F., Näätänen, R., & Ilmoniemi, R. J. (2007). Grammar processing outside the focus of attention: an MEG study. *Journal of Cognitive Neuroscience, 15*(8), 1195-1206. https://doi.org/10.1162/089892903322598148

Söderström, P., Horne, M., & Roll, M. (2016). Stem tones pre-activate suffixes in the brain. *Journal of Psycholinguistic Research*, *46*, 271-280. https://doi.org/10.1007/s10936-016-9434-2

Söderström, P., Horne, M., Frid, J., & Roll, M. (2016). Pre-activation negativity (PrAN) in brain potentials to unfolding words. *Frontiers in Human Neuroscience, 10*(512), 1-11. https://doi.org/10.3389/fnhum.2016.00512

Söderström, P., Roll, M., & Horne, M. (2012). Processing morphologically conditioned word accents. *The Mental Lexicon, 7*(1), 77-89. https://doi.org/10.1075/ml.7.1.04soe

Sorace, A. (1985). Metalingistic knowledge and language use in acquisition-poor environments. *Applied Linguistics, 6*(3), 239-254. https://doi.org/10.1093/applin/6.3.239

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 & Language, 120*, 135-162. https://doi.org/10.1016/j.bandl.2011.07.001

Tanner, D., Morgan-Short, K., & Luck, S. J. (2015). How inappropriate high-pass filters can produce artifactual effects and incorrect conclusions in ERP studies of language and cognition. *Psychophysiology, 52*, 997-1009. https://doi.org/10.1111/psyp.12437

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

van Lancker, D. (1980). Cerebral lateralization of pitch cues in the linguistic signal. *Paper in Linguistics, 13*(2). https://doi.org/10.1080/08351818009370498

Wicha, N. Y., Moreno, E. M., & Kutas, M. (2003). Expecting gender: An event related brain potential study on the role of grammatical gender in comprehending a line drawing within a written sentence in Spanish. *Cortex, 39*(3), 483-508. https://doi.org/10.1016/S0010-9452(08)70260-0

Wicha, N. Y., Moreno, E. M., & Kutas, M. (2004). Anticipating words and their gender: an event-related brain potential study of semantic integration, gender expectancy, and gender agreement in Spanish sentence reading. *Journal of Cognitive Neuroscience, 16*(7), 1272-1288. https://doi.org/10.1162/0898929041920487

Wiese, R. (1996). Phonological versus morphological rules: On German umlaut and ablaut. *Journal of Linguistics, 32*(1), 113-115. https://doi.org/10.1017/S0022226700000785

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. (2002). *Tone.* Cambridge University Press. https://doi.org/10.1017/CBO9781139164559

Yu, K., Li, L., Chen, Y., Zhou, Y., Wang, R., Zhang, Y., & Li, P. (2019). Effects of native language experience on Mandarin lexical tone processing in proficient second language learners. *Psychophysiology, 56*, Article e13448. https://doi.org/10.1111/psyp.13448

Yu, K., Wang, R., Li, L., & Li, P. (2014). Processing of acoustic and phonological information of lexical tones in Mandarin Chinese revealed by mismatch negativity. *Frontiers in Human Neuroscience, 8*(729), 1-9. https://doi.org/10.3389/fnhum.2014.00729

Yu, L., & Odlin, T. (2016). *New Perspectives on Transfer in Second Language Learning.* Multilingual Matters.

Yue, J., Bastiaanse, R., & Alter, K. (2014). Cortical plasticiy induced by rapid Hebbian learning of novel tonal word-forms: Evidence from mismatch negativity. *Brain & Language, 139*, 10-22. https://doi.org/10.1016/j.bandl.2014.09.007

Appendices

Appendix A: Overview of Swedish language background for non-tonal participants

Learner ID*	Weeks in Sweden	Weeks of Swedish	Intensity (hours/week)	Self-rated listening proficiency in Swedish**
g01	8	20	2	novice
g02	24	24	2	B1
g03	7	5	6	novice
g05	11	11	0.5	A1
g06	12	10	5	B1
g07	41	0	0	A1
g08	15	2	N/A	none
g09	12	2	6	novice
g10	14	2	5	novice
g11	14	2	N/A	A1
g12	14	14	8	B1
g13	15	15	10	B1
g14	16	6	4	A2
g15	16	6	4	novice
g16	17	8	2	A1
g17	24	24	5	B1
g18	25	4	0.2	novice
g19	21	4	4	A2
g20	42	0	0	A2
g21	53	12	3	B1
g22	78	12	1.5	A2
g23	30	22	8	A2
g24	16	16	6	A2/B1

* The learner that was excluded from the analysis (g04) is not listed in the table, for paper II, learner g05 and g08 were excluded due to low accuracy at the mismatch detection task of the artificial learning study

** none =not even single words; novice=single words; A1=familiar expressions and phrases; A2=sentences and common expressions; B1=main points of common conversations

Paper I

Journal of Neurolinguistics 45 (2018) 60–78

Contents lists available at ScienceDirect

Journal of Neurolinguistics

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

Neural processing of morphosyntactic tonal cues in second-language learners

Sabine Gosselke Berthelsen [a,*], Merle Horne [a], K. Jonas Brännström [b], Yury Shtyrov [c,d], Mikael Roll [a]

[a] Department of Linguistics and Phonetics, Lund University, Sweden
[b] Clinical Sciences Lund, Division of Logopedics, Phoniatrics and Audiology, Lund University, Sweden
[c] Center of Functionally Integrative Neuroscience, Aarhus University, Denmark
[d] University of St. Petersburg, St. Petersburg, Russian Federation

A R T I C L E I N F O

Article history:
Received 23 November 2016
Received in revised form 21 August 2017
Accepted 13 September 2017
Available online 6 October 2017

Keywords:
Second language acquisition
Prediction
Morphosyntax
Word accents
Pre-activation negativity
ERP

A B S T R A C T

The morphosyntactic nature of word accents in Swedish makes them a perfect candidate for the study of predictive processing in language. The association of word stem accents with upcoming suffixes allows native listeners to pre-activate a word's potential ending and thereby facilitate speech processing. Unlike native speakers, second language learners are known to be less able to use prediction in their L2s. This is presumably due in particular to competing information from the learners' L1 and a poorer exposure to the relevant L2 information. Swedish word accents, however, are abundant in the input and rare cross-linguistically, making them ideal for studying the implicit acquisition of linguistic prediction in beginner L2 learners. We therefore recorded learners' electrophysiological brain responses to Swedish word accents and compared them to those of native speakers. In the native speaker group, a pronounced suffix-related PrAN (pre-activation negativity), N400 and a P600-like late positivity indicate predictive processing. The learners, however, only produced a late (400–600 ms) centrally distributed negativity for word accent processing, remarkably similar to the deflection for pure pitch height differences found in the same subject group. Crucially, correlation analysis indicated that this negativity increased (at right-lateral electrode sites) for learners with increased level of Swedish proficiency. We conclude that, to allow L2 tone-suffix association and to enable its predictive capacity, the acquisition of Swedish word accents and their predictive properties might first involve dissociation of word tones from the default L1 tonal patterns as well as sensitisation to pitch height differences.

1. Introduction

Understanding spoken language is a complex and resource-demanding task. The work-load is even higher if the language to be processed is non-native to the listener (Hasegawa, Carpenter, & Just, 2002). The present study set out to investigate whether beginner learners of a language can implicitly learn and make use of predictive prosodic cues that do not exist in their native language in order to facilitate the processing of upcoming morphological information.

* Corresponding author. 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.jneuroling.2017.09.001

1.1. Prediction in language processing

The idea that listeners use predictability in constrained contexts to facilitate the processing of upcoming information has influenced language research since the 1970s (e.g. DeJong, 1979; Fischler & Bloom, 1979). Accelerated lexical decision times for correctly pre-activated information in constrained sentences has been seen as proof thereof (Fischler & Bloom, 1979). However, it has remained difficult to show that these accelerated lexical decision times result from a word's higher contextual predictability rather than eased integration of the word into the sentential context as traditional bottom-up language processing theories would claim (e.g. Jackendoff, 2002; Norris, 1994). More recently, however, the use of event-related potentials (ERPs) to directly sample the brain's activity dynamics during language tasks has produced a growing body of evidence for the psychological reality of prediction in language comprehension and production.

1.1.1. Benefits and consequences of predictive processing

In a study on compound nouns, Koester, Gunter, Wagner, and Friederici (2004) observed a left anterior negativity (LAN) for both the first and the second constituent in the compound if they were mismatched in gender with a preceding determiner. The LAN for the first noun in compounds - where a match in gender is irrelevant - suggests that listeners automatically make predictions about the gender of the upcoming noun when encountering a gender-specific determiner. Placing noun phrases into sentential contexts, other studies have looked at the processing of articles which by themselves carry little differential semantic weight. By investigating articles, influences of integration could be minimalised and effects of prediction could more easily be factored out. It has been shown that listeners produce a higher N400 when encountering an article of an unexpected gender (Wicha, Moreno, & Kutas, 2003) or with an unexpected phonological form (DeLong, Urbach, & Kutas, 2005) in contextually constrained sentences. Both the unexpected article and the following unexpected noun have been seen to elicit a greater N400, which, furthermore, is observed to correlate with the noun's offline cloze probability (DeLong et al., 2005). In contextually constrained contexts, a reduced N400 to correctly predicted words can thus be said to function as an index of the processing benefits of prediction.

The reason why these studies found differences in the N400 while Koester et al. (2004) found a LAN presumably lies in the fact that the latter looked at noun phrases in isolation and hence had a more morphosyntactic focus, while Wicha et al. (2003) and DeLong et al. (2005) placed nouns into sentential contexts and tasked participants with comprehension or memory tasks which focus more on the lexical-semantic content. The LAN and the N400 have been argued to be very similar effects that are slightly more influenced by morphosyntax and semantics, respectively (Molinaro, Barber, Caffarra, & Carreiras, 2014). This is corroborated by studies on predictive tones (i.e. word accents, see section 1.1.2) in Swedish, which have found a LAN for failed predictions in the absence of semantic information, i.e. using pseudowords (Söderström, Horne, & Roll, 2016b). Unpredicted suffixes in real words have rather produced an N400-like increased negativity (Roll, 2015).

Akin to the early negativities, a late positivity or P600 has frequently been found to follow the LAN or N400 in predictive contexts (e.g. Federmeier, Wlotko, De Ochoa-Dewald, & Kutas, 2007; Hahne & Friederici, 1999; Koester et al., 2004). The P600 has been argued to show consequences of failed prediction (Federmeier et al., 2007; Kutas, DeLong, & Smith, 2011) i.e. the need of repair or update of a context model originating from the unexpected resolution of a wrongly predicted sentence continuation. The anterior-posterior distribution of the late positivity in this context varies between studies; while explanations for this variability diverge somewhat, researchers seem to agree that the frontal effect stems from the resolution of ambiguity for unexpected words in contextually constrained sentences (e.g. Kaan & Swaab, 2003; Kuperberg & Jaeger, 2016; Kuperberg, 2013; van Pretten & Luka, 2012). This effect would be evoked for instance when ending the sentence "His skin was red from spending the day at the" with the unexpected, but plausible ending "farm" rather than the contextually expected "beach" (cf. Federmeier et al., 2007). An unexpected but fully plausible ending does not as such call for repair. The frontal component thus rather signals an ambiguity resolution (Kaan & Swaab, 2003). The posterior distribution of the positivity, on the other hand, is argued to stem from an implausible/incongruent or morphosyntactically erroneous continuation. This is seen for instance in verbs or adjectives that mismatch the obligatory number agreement prediction generated by preceding noun phrases (Kaan & Swaab, 2003; Roll, Gosselke, Lindgren, & Horne, 2013), suffixes that mismatch the number prediction generated by prosodic cues (Roll, 2015; Söderström et al., 2016b) or verbs that mismatch the event structure prediction generated by thematic roles (Kaan & Swaab, 2003; Kuperberg, 2013), i.e. incongruent sentence completions (van Pretten & Luka, 2012). The ERP response in this context indicates the need for revision or repair.

1.1.2. Online predictive processing: Swedish word accents and pre-activated endings

By highlighting the connection between ERP events and related preceding conditioning contexts, the above studies have provided strong evidence for the existence of predictive strategies in language processing. However, they do not tap into the pre-activation process itself. Predictions are most often gradually built up throughout sentence processing and there is not usually a clear cut-off point at which a prediction is formed. There are, however, exceptions. The Swedish language has a prosodic system at the word-level which modulates the predictability of an immediately upcoming suffix and thus pinpoints a precise moment in speech processing where predictions can be made. This allows for the study of ongoing pre-activation. Previous research in this area has shown that native listeners, upon hearing prosodic cues on word stems, are able to pre-activate the upcoming word ending (Roll, 2015; Roll et al., 2015; Söderström, Horne, Frid, & Roll, 2016; Söderström et al., 2016b). To fully follow the argument, it is important to understand the Swedish prosodic system in more detail:

In Swedish, there are two so-called ‘word accents’, accent 1 (low tone) and accent 2 (high tone) which are realised on the word stem, i.e. before the suffix onset. The definite singular suffix *-en*, for instance, is associated with accent 1 and will therefore co-occur with a low tone on a preceding nominal stem like *båt* in *båten* (boat-the, cf. Fig. 1). The indefinite plural suffix *-ar*, in contrast, is associated with accent 2, which is realised as a high tone on the preceding nominal stem (cf. high tone in Fig. 1) (Riad, 1998; Rischel, 1963). The stem tones can therefore be used as cues in speech processing for pre-activating upcoming endings, thus establishing predictability and easing processability (Roll, Horne, & Lindgren, 2010; Roll, Söderström, & Horne, 2013; Söderström, Roll, & Horne, 2012). In Swedish, accent 1 has been shown to be a better predictor than accent 2, since it can be followed by a limited set of suffixes (e.g. definite singular *-en* and present tense *-er*). Accent 2, while also cueing suffixes (e.g. indefinite plural *-ar* and past tense *-te*), is additionally associated with productive compounds and thus can be assumed to pre-activate a much larger set of possible continuations for a given word stem, which reduces its predictive value (Söderström et al., 2016). Unexpected suffixes due to a mismatch in word stem tone and following suffixes, give rise to well-established predictive processing consequences. Behaviourally, this is reflected in increased response times to word accents with mismatched suffixes, where responses are considerably delayed due to failed pre-activation (Söderström et al., 2012). Neurophysiological studies have found that mismatched and thus unexpected suffixes result in the ERPs typically seen for failed predictions, such as LAN, N400 and P600 (Roll, 2015; Roll, Söderström et al., 2010, 2013, 2015; Söderström et al., 2016b).

In addition to studying known ERP effects for contextually unexpected linguistic items, Swedish word accents make it possible to investigate the online implementation of predictive processes and the pre-activation of upcoming syllables. In this regard, an early negative component has consistently been observed at 136–280 ms after onset of the stem tone. The stem tone pinpoints a specific moment in time at which predictions about the upcoming suffix can first be made. As the negativity occurs during the processing of the word stem and thus before suffix onset and before reliable grammaticality judgements can be made, it must be regarded as being triggered by the stem tone, indexing its function in pre-activating possible upcoming continuations. Due to this cognitive content, this negative component has been labelled ‘pre-activation negativity’ (PrAN; Roll, 2015; Roll et al., 2015; Söderström, Horne, & Roll, 2016a, 2016b).

In support of the claim that the PrAN indexes pre-activation of upcoming suffixes, a negative correlation has been observed between the number of word completions a given word stem fragment can be associated with and the amplitude of the PrAN (Söderström et al., 2016). Stem fragments with fewer possible continuations (i.e. less lexical competition) have been shown to have greater predictive value, resulting in a larger pre-activation negativity. The restricted set of endings for accent 1 stems thus produces a comparably large pre-activation negativity in native speakers of Swedish, which provides strong evidence for a solid association between stem tone and upcoming suffix in Swedish.

1.2. Second language acquisition: prediction and morpheme learning dynamics

In view of how important language processing properties (such as prediction) work in speakers' native languages, the question arises as to what extent they are activated in non-native speakers' language processing. Can language learners use predictions in their second language much like they do in their first one? If continuous prediction and pre-activation of potential upcoming information is a key feature in language processing, it stands to reason that it would be an important

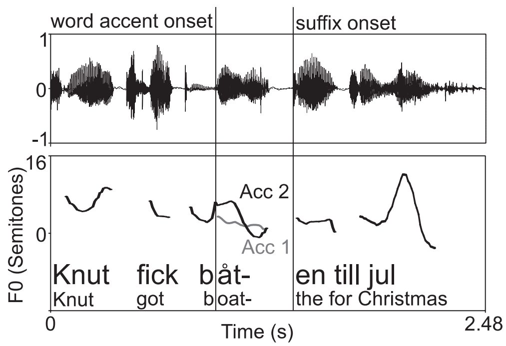


Fig. 1. Example of a stimulus sentence with acoustic waveform and fundamental frequency (F0). The two F0 contours on *båt* show the valid low accent 1 (grey) and invalid high accent 2 (black) for the definite singular ending *-en*.

feature in any language, native or not. It may even be possible that general prediction mechanisms facilitate acquisition in that they function as processing aids, thereby releasing cognitive resources for sentence processing.

1.2.1. Prediction in second language learners

Previous research has yielded mixed results as regards the prototypical behavioural and ERP effects to prediction mismatch in non-native speakers. Near-null results for L2 prediction have been reported by Guillelmon and Grosjean (2001), who, using behavioural measures, found that the identification of nouns was faster when they were followed by gender marked articles. This effect was only seen for native speakers and in reduced form for early bilinguals. Highly proficient learners who acquired the L2 at a later age did not show any behavioural facilitation effects of gender marked articles for the identification of nouns. A more predictive yet non-native-like type of processing in high proficiency late bilinguals was found in Martin et al. (2013), who reported that second language learners, much like native speakers, produced an N400 for unexpected nouns. However, unlike L1s, they failed to produce an N400 for unexpected articles. Finally, Foucart, Martin, Moreno, and Costa (2014) found identical effects for native speakers and second language learners: an N400 effect for articles and nouns and a late anterior positivity for nouns. Whether or not learners produce a PrAN has not yet been studied, but behavioural results, previously correlated with the PrAN, suggest that at least advanced learners might do so (Schremm, Söderström, Horne, & Roll, 2016).

A plausible explanation for the mixed results for predictive processing in second language learners is offered by Kaan (2014) who attributes prediction differences in L2s to several underlying differences, most importantly frequency information and competing information. She claims that second language learners are in general exposed to less input than native speakers and that this input is of different quality. The strength of predictions they generate is therefore weaker than for L1s. Additionally, learners may have competing information from their L1 that may be activated automatically even in non-native speech context, something which could affect the acquisition of pre-activation strategies in their L2. If the neural connections between different kinds of potentially predictable linguistic information are not strong and stable, pre-activation will be hampered, and hence the lack of effects for failed predictions. Consequently, learners are likely to only acquire pre-activation when the triggers (1) are highly frequent in the linguistic input and (2) do not clash with native language features. Swedish word accents may hence be ideal for investigating prediction and pre-activation in second language learners. They are realised on every word in Swedish making them abundant in the linguistic input. They thus constitute a good candidate for statistical learning and allow formulation of strong predictions. Furthermore, word accents are virtually exclusive to the Scandinavian languages and should therefore not have to compete with the learners' native language processing strategies.[1] At the same time, it is important to note that their knowledge appears to be acquired implicitly and is not taught as part of Swedish tuition; in fact, most native speakers are unaware of their own use of this feature (Roll, Söderström, et al., 2013).

1.2.2. Morpheme learning dynamics in second language learners

Language acquisition research has shown that novel morphemes in a speaker's native language can be acquired within mere minutes of contact. The learner's brain responses adapt rapidly and start resembling those of real morphemes whether taught explicitly (Leminen, Kimppa, Lehtonen, Mäkelä, & Shtyrov, 2016), by means of context (Mestres-Missé, Rodriguez-Fornells, & Münte, 2007) or merely presented repeatedly outside of language contexts, i.e. acquired through implicit, statistical learning (Kimppa, Kujala, Leminen, Vainio, & Shtyrov, 2015; Kimppa, Kujala, & Shtyrov, 2016; Shtyrov, 2011; Shtyrov, Nikulin, & Pulvermüller, 2010). In contrast to native-language-like morphemes, ERP responses to repeatedly presented novel words with non-native phonology do not become real-word-like as quickly, suggesting that implicit acquisition is less rapid for morphemes in learners' non-native languages (Kimppa et al., 2015, 2016). While the acquisition of non-native morphemes and phonology is likely to be delayed compared to native language equivalents, it can still be relatively fast. In an explicit learning paradigm, for instance, McLaughlin, Osterhout, and Kim (2004) showed that 14 h of classroom tuition were sufficient for learners' brain responses to correctly differentiate between novel words and pseudowords from the still fairly unknown language. Behavioural evidence suggested that this phonotactic awareness in fact emerges even faster and also during implicit learning. Only seven minutes of exposure to visually emphasised, restricted but natural speech in a foreign language are sufficient for the listeners to statistically extract both implicit phonotactic rules and word meaning (Gullberg, Roberts, & Dimroth, 2012; Gullberg, Roberts, Dimroth, Veroude, & Indefrey, 2010).

In this light, it is hardly surprising that Schremm et al. (2016) found that behavioural responses to word accents in relatively advanced, upper intermediate learners of Swedish resembled those of native speakers, albeit less pronounced. Interestingly, although the learners were not yet highly proficient and the studied feature (i.e. word accents) not taught explicitly, their behavioural measures indicate that they had a good grasp of Swedish phonology, morphology and prosody. The difference in response times to accent 1 and accent 2 suggests that they had extracted and acquired the predictive features of Swedish word accents and were in fact able to pre-activate suffixes in a manner similar to native speakers.

Considering how fast even non-native language acquisition may progress, the question arises whether even beginner learners might be able to extract the tonal features associated with word accents and use them predictively in accessing

[1] The learners in the present study had German as L1. Although German does not have word tones, it does have pitch accents that are associated with pragmatic interpretations related to information structure and speaker attitude (Grice et al., 2005). Consequently, it could be expected that the German L2 learners might initially interpret Swedish word accents as pitch accents associated with some kind of pragmatic meaning.

Table 1
Average stimulus lengths and standard deviations.

	NP + V	Target word				PP
		Accent 1		Accent 2		
		Stem	Suffix	Stem	Suffix	
Example	*Knut fick* Knut got	*båt-* boat	*-en* -DEF.SG	*båt-* boat	*-ar* -PL	*till jul* for Christmas
Mean, ms	806	547	241	544	241	873
±SD, ms	59	59	24	68	24	68

morphosyntactic information. In the current study, we therefore look at a beginner group of Swedish L2 learners in order to determine whether there is evidence of a predictive use of word accents in speech perception and understanding or, alternatively, whether there is evidence of any stepping stones that may pave the way for predictive processing at more advanced levels.

2. Method

2.1. Participants

Twenty-three native speakers of Swedish (L1s; 13 females) and twenty-three native German-speaking learners of Swedish (L2s; 12 females) participated in this study.[2] None of the German participants were speakers of pitch accent dialects of German or had any extended exposure to such dialects. All were right-handed as assessed by the revised Edinburgh Handedness Inventory (Williams, 2013) and had normal hearing defined as pure-tone hearing thresholds ≤20 dB Hearing Level (ISO, 2004) for frequencies 250, 500, 1000, 2000, 4000, and 8000 Hz. The participant groups were matched for age (mean 23.8 years, SD = 2.1 years), socioeconomic status as rated using Hollingshead's (1975) index, and working memory span as tested by the automated operation span task (Unsworth, Heitz, Schrock, & Engle, 2005).

The self-rated proficiency in spoken Swedish comprehension in the L2 group (German native speakers) showed an average level of beginner learner (=being able to understand everyday expressions and very basic phrases, i.e. level A1 of the Common European Framework of Reference), ranging from pre-beginners (=being able to make out single words in a conversation at most) to early intermediates (=being able to understand main points of conversations in familiar contexts, i.e. level B1). All participants in the L2 group were exchange students at Lund University and had lived in Sweden for an average of 160 days (SD = 114 days). The continuous span in proficiency and residency in Sweden was suitable for correlation analyses to assess the learners' gradual ongoing development at this level.

Due to living abroad at the time of testing, the L2 group reported a temporary increase in the use of non-native languages in relation to Swedes (63.05% vs. 30.91%). Yet, importantly, the groups were matched in the number of languages spoken at above elementary proficiency and no participant had previously learned any tone languages other than Swedish. During background testing, both groups displayed an equal ability to detect non-linguistic pitch differences and durational variations in segments (vowel length).

The experiments were conducted in agreement with the ethical guidelines for experiments in the Declaration of Helsinki.

2.2. Stimuli

2.2.1. Word accents

The stimuli for the word accent part of the study were identical to those in Roll et al. (2015). The experiment hence contained 30 sentences with target words in 4 systematically manipulated conditions (singular/plural X valid/invalid; Table 1). The target words were embedded into carrier sentences of the type 'NP got <u>target</u> PP':

Knut	*fick*	*<u>båten</u>*	/	*<u>båtar</u>*	*till*	*jul.*
Knut	got	<u>boat-DEF.SG</u>	<u>/</u>	<u>boat-PL</u>	for	Christmas.
'Knut got <u>the boat/boats</u> for Christmas'						

The final prepositional phrase was focused so that test words could be associated with non-focal word accents (Bruce, 1977). Voiceless stops were chosen to border on the target word both to the left (*fick*) and to the right (*till*). Since these

[2] Twenty-four participants were originally recruited in each group. One Swedish participant experienced discomfort with the EEG equipment and consequently asked to quit the study. One German participant had rudimentary knowledge of a tone language and was hence excluded from the analysis.

stops are associated with silent closure periods, a clear-cut separation of the target word from surrounding words was possible. Similarly, the first syllable of the target word ended in a voiceless stop to facilitate determining suffix onset position. For stimulus sentence lengths and standard deviations see Table 1.

The same sets of carrier phrases occurred in each condition; equal shares were taken from the singular and plural target word recordings. Stems and suffixes were cross-spliced and used in both valid and invalid conditions. The accent 1 tone had a duration of 137 ms, SD = 15 and fell from 2.71 semitones at vowel onset, SD = 0.86, to 1.30 semitones at vowel offset, SD = 0.52. The accent 2 tone was marginally shorter (M = 117 ms, SD = 27) but considerably steeper, falling from 6.76 semitones at vowel onset, SD = 1.53, to 3.08 semitones at offset, SD = 2.25. The time between stem tone onset and suffix onset was 426 ms (SD = 63, range = 348–587). All sentences were normalised to the same power, measured as root-mean square (RMS) amplitude. The resulting 120 sentences were pseudorandomised and divided into two blocks. Every block contained additionally 20 silent trials to prevent predictability and keep participants alert.

2.2.2. Piano tones

To assess test persons' electrophysiological auditory responses to differences in tone heights unconfounded by linguistic context or Swedish-language background, they participated in a test involving pitch discrimination. To this purpose, piano tunes were used as stimuli. 20 consecutive piano tones were recorded, ranging from $B\flat_3$ to F_5. The tones were 500 ms long each and normalised to the same power. Tones were combined into pairs where the difference between tone A and tone B was between one and eight semitones. There were eight trials for each semitone distance; 64 trials in total. There was an equal number of high-low and low-high tone pairs.

2.3. Experimental procedure

Participants were seated at a distance of 1 m from a computer screen. A response box was placed on a table in front of them and they there were asked to keep their index fingers on the response buttons throughout the experiment. All 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.

2.3.1. Word accents

For the word accent part of the study, participants listened to the Swedish sentence stimuli (120 trials in total). In each trial, one stimulus sentence was presented auditorily and participants executed a simple sentence boundary decision task in which they hit a button on the response box, alternatingly with their right and left index fingers, as soon as they perceived the sentence to be over. Sound presentation and response had a fixed total duration of 3.5 s. Depending on sentence length, the time allowed for responding was 0.88–1.27 s. Each trial was concluded by a 500 ms silent pause, and thus the total onset-to-onset stimulus onset asynchrony (SOA) was 4 s. A fixation cross remained on the screen for the entire duration of the trial, including the pause. The presentation automatically moved on to the next sentence, even if a response was not made. No feedback was given.

The chosen sentence boundary detection task was feasible even for the L2 group with little to no comprehension of the sentences' semantic content, due to the learners' background: German and Swedish intonation is similar as regards neutral declarative sentence boundary tones (Gibbon, 1998; Gårding, 1998). As the stimuli in this study consisted entirely of declarative sentences with focus on a sentence-final prepositional phrase, we assume that the learner group could detect sentence boundaries even if they failed to understand (some of) the sentence meaning. This is supported by the behavioural results that indicated similar accuracy in both groups (see the Results section below).

2.3.2. Piano tones

For the pitch discrimination part of the study, participants listened to pairs of piano tones. After a 750 ms fixation cross, tone A was presented auditorily for 500 ms, followed by a 750-ms fixation cross and a 500-ms presentation of tone B. A 400-ms fixation cross separated the tone pair from a response screen where subjects were prompted to indicate which tone was higher. They did so by pressing the left button on the response box for tone A and the right button for tone B with their left and right index finger, respectively.

2.4. Electroencephalography (EEG)

Throughout the experiment, the participants' brain activity was continuously 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). Horizontal and vertical bipolar electrooculogram channels (EOG) were added to record eye movements. 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.

Offline, the data was re-referenced to average reference and filtered with a 0.01 Hz high pass filter and a 30 Hz low pass filter. 1200-ms long ERP epochs including a 200-ms baseline were extracted both at word accent onset and at suffix onset for word accents, and at tone A and tone B onset for piano tones. Eye artefacts were removed with the help of independent component analysis (ICA) (Jung et al., 2000) and all epochs still exceeding ±100 μV were discarded.

2.5. Statistical analysis

2.5.1. Word accents

For the behavioural data, mean response times (RTs) for valid (i.e. not premature) responses were submitted to a repeated measures Analyses of Variance (rmANOVA) with within-subject factors 'Suffix' (*-en* vs. *-ar*), 'Validity' (valid vs. invalid) and the between-subject factor 'Group' (L1 vs. L2). All statistical analyses were performed in IBM SPSS Statistics 22 (International Business Machines Corp., Armonk, NY, United States) and Greenhouse–Geisser correction was used when applicable.

The ERP data for word accent onset and suffix onset were analysed separately. For word accent onset, two different analysis time windows were selected. The first one corresponds strictly to the previously established time window for the PrAN, i.e. 136–280 ms, originally chosen on the basis of global root mean square (gRMS) peaks and their correlations with blood-oxygen-level dependent (BOLD) effects (Roll et al., 2015). The second time window was based on a peak observed in the L2 group's ERP data. Statistical analyses using a time window sliding in 20 ms steps revealed 400–600 ms as the maximal time window in which the observed effect was significant. Similarly, visual inspection and statistically optimised time windows led to the two analysis windows for the suffix onset data: 235–415 ms and 550–680 ms.

Mean amplitudes for the word accent onset response latencies were submitted to an rmANOVA with the experimental factor 'Word Accent' (accent 1 vs. accent 2) and the topographical distribution factors 'Laterality' (left, mid, right) and 'Posteriority' (anterior, central, posterior); corresponding to electrode groups as follows: 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. For latencies above 350 ms after word accent onset, the experimental factor 'Validity' (valid vs. invalid) was added to the rmANOVA. This was done due to the fact that the suffix emerges in the input material from this point onward which might influence the word accent responses. In a parallel manner, mean amplitudes for the two suffix onset latencies were submitted to an rmANOVA with the experimental factors 'Suffix' (*-en* vs. *-ar*) and 'Validity' (valid vs. invalid) and the distributional factors 'Laterality' (left, mid, right) and 'Posteriority' (anterior, central, posterior).

To investigate whether the amplitudes of L2 learners' effects were related to their individual degree of contact with the Swedish language, the L2s' mean amplitude for each laterality factor and the behavioural variables 'Level of Swedish', 'Hours of Tuition', and 'Weeks in Sweden' were submitted to a 2-tailed Pearson correlation. Bonferroni corrections for multiple comparisons were applied.

2.5.2. Piano tones

With regard to the behavioural measures, an independent-samples *t*-test was conducted to compare response times and accuracy rating for the L1 and the L2 participants. For the ERP data, visual inspection revealed a negativity which in both latency and distribution resembled the one found in the L2s' word accent onset data. In order to directly compare the two, the same time window (400–600 ms) and relevant topographical factor ('Laterality') were used. Mean amplitudes were submitted to an rmANOVA with the experimental factor 'Tone Height' (low vs. high) and the topographical factor 'Laterality' (left, mid, right).

3. Results

3.1. Word accents

3.1.1. Behaviour

Both participant groups produced a number of premature responses before the end of the sentence ($M_{NS} = 25.69\%$, $SD = 19.26$; $M_{L2} = 25.07\%$, $SD = 20.19$). These responses were excluded from the analysis.

Response times to sentence boundaries revealed a Suffix*Group interaction, $F(1,44) = 4.51$, $p = 0.039$, which unfolded into a main effect of Suffix in the native speaker group only, $F(1,22) = 5.03$, $p = 0.035$. Native speakers' response times for sentences with the accent 2-associated suffix *-ar* ($M = 224$, $SD = 93$) were significantly slower than RTs for sentences with the accent 1-associated suffix *-en* ($M = 214$, $SD = 93$). There were no statistically significant differences in response times for the second language learner group ($M = 219$, $SD = 92$).

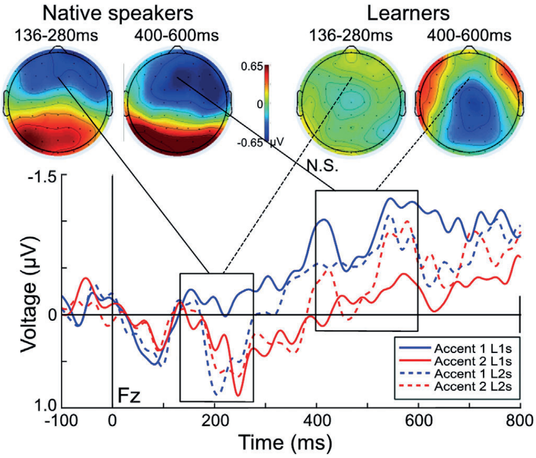


Fig. 2. Above: Topographical distributions of the early and late accent 1 – accent 2 effects (136–280 ms and 400–600 ms) for native speakers (continuous lines) and learners (dotted lines). Below: ERPs at mid-anterior electrode Fz for the two groups.

3.1.2. *Electrophysiology*

3.1.2.1. Word accent onset. For the native speakers, accent 1 yielded a negativity in both the PrAN time window (136–280 ms) as well as at 400–600 ms after tone onset. In the earlier time window, this led to a Word Accent*Posteriority interaction, $F(2,44) = 9.20$, $p = 0.004$, which was due to an increased negativity ($F(1,22) = 8.83$, $p = 0.007$) for accent 1 at anterior but not posterior electrode sites. Similarly, at 400–600 ms, a Word Accent*Posteriority interaction, $F(2,88) = 9.48$, $p = 0.004$, revealed a significant negativity for accent 1 at anterior sites, $F(1,22) = 11.46$, $p = 0.003$ (cf. Fig. 2). There was no significant interaction with the suffix factor Validity.

For the L2s, accent 1 produced a negativity at 400–600 ms. Its analysis showed a Word Accent*Laterality interaction, $F(2,44) = 3.30$, $p = 0.049$, and a negativity for low accent 1 at mid electrode sites, $F(1,22) = 6.13$, $p = 0.021$. Further analysis of the L2 group indicated that there was a significant correlation between Level of Swedish and negativity amplitude at right-hemispheric electrode sites (cf. Fig. 3); $r = -0.511$, $p = 0.039$ (Bonferroni-corrected). No other factors correlated with the effect.

3.1.2.2. Suffix onset. The L1 group produced a negativity for invalid suffixes specific to mid lateral electrode sites, which was reflected in a near-significant Validity*Laterality interaction at 235–415 ms after suffix onset, $F(2,44) = 3.41$, $p = 0.051$, and resolved in a significant effect for Validity at mid electrodes, $F(1,22) = 6.84$, $p = 0.016$. Similarly, a broadly distributed

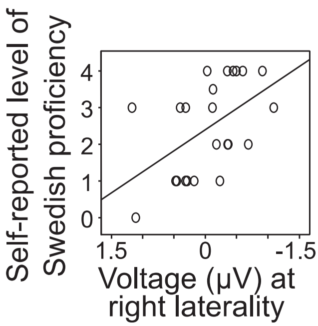


Fig. 3. Scatterplot for the correlation of negativity at right lateral electrodes and self-reported level of Swedish proficiency in the L2 group: 0 = lowest (~pre-beginner) to 4 = highest (~intermediate learner (B1)).

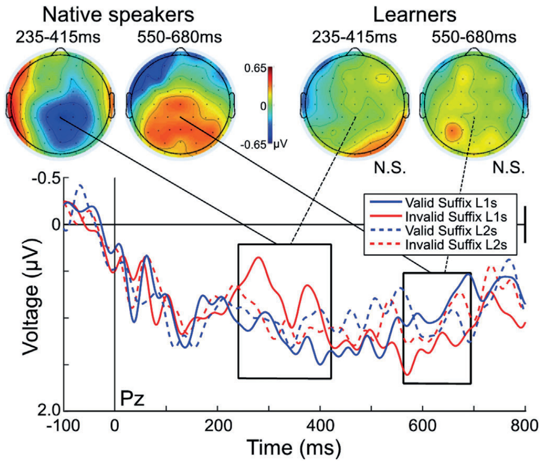


Fig. 4. Above: Topographical distributions of the early and late validity effects (235–415 ms and 550–680 ms) for native speakers (continuous lines) and learners (dotted lines). Below: ERPs at mid-posterior electrode Pz for the two groups.

positivity for invalid suffixes was marginally significant at 550–680 ms, $F(1,22) = 3.77$, $p = 0.065$ (cf. Fig. 4). There were no effects in the L2 group for the ERPs time-locked to the suffix onset.

3.2. Piano tones

3.2.1. Behaviour

Response times and accuracy ratings for the two participant groups were compared in independent-samples t-tests. Accuracy of the tone height assessments was high overall ($M_{L1} = 91.24\%$, $SD = 9.0$; $M_{L2} = 93.75\%$, $SD = 6.2$) and did not differ significantly between groups; $t(44) = -1.08$, $p = 0.285$. To normalise the distribution of participants' average response times

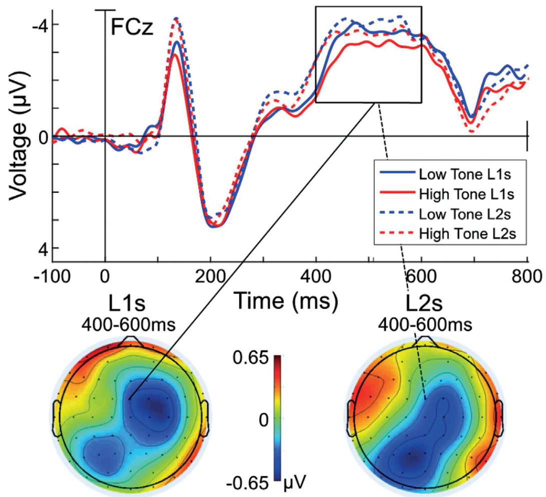


Fig. 5. Above: ERPs at mid-anterior electrode FCz for contrasting low (blue) and high tones (red) for native speakers (continuous lines) and learners (dotted lines). ERPs are time-locked to tone onset. Below: Topographical distribution of the tone height related negativity (low-high) from 400 to 600 ms for L1s (left) and L2s (right).

(original RTs: $M_{L1} = 835$ ms, $SD = 511$; $M_{L2} = 696$ ms, $SD = 277$), we used log-transformed response times for the statistical analysis. They revealed no significant differences between groups ($t(44) = 0.77$, $p = 0.445$).

3.2.2. Electrophysiology

Listening to high-low and low-high pairs of piano tones produced a negativity for the lower tones in the pair at 400–600 ms after tone onset. For the L1 group, this manifested as a main effect for Tone Height, $F(1,22) = 8.50$, $p = 0.008$. For the L2s, it gave rise to a Tone Height*Laterality interaction $F(2,44) = 3.68$, $p = 0.034$, and an effect of Tone Height at mid lateral electrode sites, $F(1,22) = 6.10$, $p = 0.022$ (cf. Fig. 5).

4. Discussion

The present study investigated whether adult beginner learners of a tone language can acquire and make use of predictive prosodic cues in the same way as native speakers do. Word stem tones in Swedish were used as a test case, as they implicitly pre-empt the morphological information delivered by word-final suffixes, which makes it possible to study both ongoing pre-activation as well as the effects of a failed prediction. Neurophysiological results showed that the native speakers produced the prototypical ERP effects associated with the predictive linguistic processing of word accents: PrAN for word accent-morphology connection followed by a negativity as well as a late positivity for word accent-suffix mismatch. The beginner L2 learner group did not exhibit similar effects. Instead, they produced a negativity which seemed to be driven purely by the non-linguistic tone height difference in the word stem tones and which expanded to right-lateral electrodes with increasing level of proficiency. We will discuss the results in more detail below.

4.1. Processing of non-linguistic pitch (piano tones)

To ensure that there were no a priori differences in pitch perception between groups, behavioural and neurophysiological responses to pitch height differences in piano tones were recorded. Results showed an overall high response accuracy, no differences in response times within or between groups and a broad centrally distributed negativity at 400–600 ms for low tones in both the native speaker and the learner group. These findings constitute a coherent indicator of similar non-linguistic pitch processing in all participants. On this account, any group differences in the processing of word accents can be assumed to be linguistic in nature, i.e. related to differences in the way they are associated with morphosyntactic information, thus ruling out any explanation based on hypothetical or spurious differences in pitch perception abilities between the two groups.

4.2. Native speakers' effects for the predictive processing of word accents

To measure the participants' behavioural performance on the Swedish sentence material with systematically modulated tonal and suffixal information, we recorded response times in a sentence boundary detection task. The native speaker group had a high rate of premature behavioural responses, which can, however, be explained by the speeded nature of the task at hand. Participants were asked to hit the response key as soon as they felt a sentence to be over. Prosodic cues aside, the sentences could have been considered complete before the final prepositional phrase, hence the large percentage of false alarms. Looking at the accepted responses only, the native group showed a clear difference in response times. Namely, boundary detection was faster in sentences with the accent 1-associated definite singular suffix *-en* compared to sentences with the accent 2-associated indefinite plural suffix *-ar*. This phenomenon has been observed previously (Roll, Söderström et al., 2013) and is likely explained by differences in morphosyntactic complexity and ease of semantic integration. The plural in Swedish has a relatively high processing complexity due to a great diversity of plural morpheme variants. This complexity could slow down semantic integration during sentence processing, resulting in decelerated response times at sentence boundary for sentences with plural as opposed to singular words. This is in line with Schremm et al.'s (2016) interpretation of longer response times for past tense suffixes as compared to present tense endings. Furthermore, in agreement with previous findings, no behavioural effect was seen for validity of the word accent-suffix combination. The lack of an effect of validity on the sentence boundary decisions suggests that prosodic violations of the word accents have limited impact on sentence meaning integration. Instead, they can be seen as a local operation, supporting ongoing morphological and lexical processing at the single-word level.

4.2.1. Native speaker EEG

4.2.1.1. Online predictive processes: stem tone effects. Like the behavioural findings, the electrophysiological markers for online predictive processing in the native speaker group were in line with previous findings. The L1 participants produced a pronounced pre-activation negativity (PrAN) for word accents which was strongest for the highly predictive accent 1 (Roll, 2015; Roll, Söderström et al., 2010, 2013, 2015; Söderström et al., 2016). As mentioned in the introduction, Swedish word stems with a low accent 1 tone are better predictors in that they can be followed by a very limited set of continuations. The word stem *båt$_1$-* (boat) with a low pitched accent 1, for instance, can only be continued with the definite singular suffix *-en* (*båt$_1$-en*, the boat). The same word stem realised with a high accent 2 tone, *båt$_2$-* (boat), in contrast, can either be continued with the indefinite plural suffix *-ar* (*båt$_2$-ar*, boats) or with one of many possible nouns constituting a part of a compound, e.g. *båt$_2$-hus* (boat house), *båt$_2$-brygga* (wharf), etc. Hence, the high tone of accent 2 on a word stem is a relatively weak predictor of upcoming endings. It is this differential predictive power of Swedish word accents that has been claimed to cause variations in PrAN amplitude, thereby rendering online predictive and pre-activating processes visible.

The assumption that differences in PrAN amplitude stem from online predictive processes gains support from an attested correlation between the PrAN's amplitude and the number of continuations for individual word stem-word accent combinations (Söderström et al., 2016). Furthermore, enhanced PrAN amplitude at stem-onset has been linked to reduced semantic judgment times, in line with its interpretation as an index of prediction/pre-activation of whole-word specific information (Roll et al., 2015; Söderström et al., 2016b).

Together, these previous and current findings strongly suggest that native speakers predict and consequently pre-activate the ending of a word stem upon hearing the tonal pattern that it carries. This is greatly facilitated when there are few competitors, resulting in a relatively larger effect. It therefore stands to reason that native speakers can make use of prosodic cues on word stems to predict how a word is going to end.

In addition to the PrAN but seemingly unrelated to predictive processing, native speakers in this study also produced a negativity at 400–600 ms after word accent onset. This effect has the same timing and deflection as an effect produced by the group when listening to pure, non-linguistic tonal differences. Its distribution, however, resembles that of the PrAN, which is believed to be produced in the left inferior frontal gyrus and the planum temporale (Roll et al., 2015). The similarity (including timing) of this deflection to the pure tone differences may suggest that the participants direct attention to the tonal information itself in this time window. However, due to the notable difference in distribution, i.e. mid-centroparietal for the non-linguistic tone effect and frontocentral for the linguistic effect, the 400–600 ms negativity for word accents may be produced in the inferior frontal gyrus, much like the PrAN, signalling a more linguistically influenced type of processing. This might be due to a more linguistic engagement with the tonal information, e.g. the tone's integration with the word stem, as well as its connection to the grammatical suffix. Alternatively, the distribution of the effect may be influenced by the simultaneous occurrence of the suffix in the linguistic input. The analysis showed that suffix validity had no impact on the word accent effect and there is thus no evidence for an interaction of word accent and suffix effects. Yet, the very appearance of additional linguistic material might affect the areas involved in the processing of this late stem tone effect.

4.2.1.2. Benefits and consequences of predictive processing: suffix effects. Besides the PrAN, signalling online predictive processing, native speakers' neurophysiology also revealed the typical processing benefits and consequences for successful and failed prediction. If the suffix was matched with the preceding, pre-activating word accent, a reduced N400-like mid-centroparietal negativity for congruent trials appeared 235–415 ms after suffix onset, reflecting eased retrieval of the pre-activated suffixed word form. Due to the semantic content of the target word, possibly aided by the general, non-morphosyntactic focus in the sentence level task, the elicitation of a more semantically influenced N400 in this type of study is not surprising.

The negativity was followed by a late posterior positivity or P600 for mismatched trials. In line with previous literature, we hypothesise that the late positivity is an index of the detection of an unexpected linguistic item. The effect's posterior distribution is probably due to fact that the prediction mismatch led to a morphosyntactic violation and the need for repair.

It is interesting to note that while the negativity for unsuccessful prediction in this study appears more sensitive to the prediction's semantic content, the late positivity rather picks up on the morphosyntactic content, resulting in a N400/posterior P600 pattern. The same pattern was observed in Roll, 2015 who used very similar stimuli, but spoken in a different dialect of Swedish. Söderström et al., 2016b, on the other hand, exchanged the real target words in the stimulus material for pseudowords. This brought about both a negativity and late positivity that picked up on the morphosyntactic part of the prediction (LAN/posterior P600), as the word stem, i.e. the part of the word that carries the predictive tone, had no semantic content. This leads us to believe that the negativity in prediction, which is believed to index prediction benefits, tends to be strongly susceptible to semantic content. The late positivity, on the contrary, which is thought to reveal consequences of failed prediction, is more influenced by type of error the mismatch leads to: unexpected, plausible endings result in a frontal positivity, indexing the resolution of an ambiguity, while unexpected and implausible or even morphosyntactically incorrect endings evoke a posterior P600 that indicates the need for revision or repair.

4.2.1.3. Linguistic function of prediction processing components. Relatively shortly after the predictive cue is spelled out in the speech input, native speakers produced a PrAN which, in the case of word accents, signals the pre-activation of relevant morphosyntactic information. It stands to reason that a similar component would appear before the onset of a contextually predicted word. The different type of prediction, i.e. contextually built up rather than instant, might change the exact timing or topographical distribution of the component. Yet we believe that the presence of a component indexing the pre-activation of upcoming material should be found in any kind of predictive linguistic environment. It is interesting to see, that after the pre-activation, a later neural response (400–600 ms) appears to index a more conscious, linguistically integrative engagement with the predictive cue, as argued above.

Once the (in)correctly pre-activated material emerges in the linguistic input, two components signal different kinds of prediction-checking processes. At first, a reduced negativity (i.e. N400 or LAN) expresses eased processing and local integration of correctly predicted morphemes. In prediction studies, the most commonly found component in this context is a reduced N400 that is argued to be representative of eased semantic integration. Rarely, a LAN is found instead, signalling local morphosyntactic rather than semantic integration with the predicting context. This appears to be the case when the predictor has a strict morphosyntax focus (e.g. when noun phrases are presented in isolation, cf. Koester et al., 2004) or when it is devoid of semantic content (e.g. when pre-activating stem tones are merged with pseudowords, cf. Söderström et al., 2016b). Subsequent to the negativity, a late positivity (i.e. frontal or posterior P600) signals hindered processing and revision or repair of incorrectly predicted morphemes. Depending on the type of prediction error, the positivity has a frontal or posterior distribution and appears to index impeded integration of the incorrectly predicted constituent as well as the possible need for revision or repair.

4.3. Beginner learners' processing of word accents

The learner group, much like the native speaker group, had a considerable number of premature responses in the behavioural boundary detection task. This was somewhat unexpected as it should have been relatively easy for the German participants to detect Swedish sentence boundaries. It was thought that the low proficiency L2s would strongly rely on intonational cues as indicators of sentence boundaries, as they are unlikely to fully understand the content of the spoken stimulus sentences. The sentence-level intonation pattern for Swedish declarative sentences is very similar to that in German (Gibbon, 1998; Gårding, 1998; Hirst & Di Cristo, 1998) and reliance on intonation patterns was therefore hypothesised to result in a low miss rate. As with the native speaker group, however, eagerness to comply with the task of responding promptly might have misled the learners into premature responses. Unlike the L1 group, however, the learners did not show any complexity-based differential effect for the different suffixes, which is indicative of their low level of Swedish proficiency. They appear to not yet grasp the often unfamiliar lexical content of the sentences, a finding which is in line with their self-reported language skills.

4.3.1. Early learner EEG

4.3.1.1. Online predictive processing: stem tone effects. Turning to the neurophysiological data, the learners' ERPs differed considerably from those of the native speakers. In contrast to the L1 participants, the L2 group did not produce any PrAN-like activation in the stem tone time window. Yet, they produced a late negativity for accent 1 words (low-pitched) at 400–600 ms at mid electrodes. This effect was crucially different in distribution from the L1's effect in this time window but virtually identical in timing and topography to the L2 group's response to non-linguistic pitch differences in piano tones. This suggests that the L2s processed the word accents non-linguistically despite their being embedded in a speech context.

In the learners' native language, tonal patterns are related to intonation and have pragmatic meaning. Through mere exposure to the abundant Swedish word accents, they appear to have learned that Swedish word accents cannot be interpreted like German pitch accents and have thus dissociated them from the context of pragmatics. The fact that they process word accents in a manner similar to their processing of pure pitch differences suggests that they are not regulated linguistically at this point. Such a dissociation of word accent pitch patterns from the L1's default pragmatic interpretation would seem to be a reasonable first step in their acquisition process.

Interestingly, we observed a subtle, yet significant variation in the beginner learners' pitch-related 400–600 ms negativity which was correlated with Swedish proficiency. The higher a participant's self-reported level of Swedish proficiency was, the greater was their negativity for word accents at right-lateral electrodes in addition to the general mid-distributed negativity. Although slightly later in timing, we suggest that this right-lateral effect could be related to an effect found for unexpected, out-of-chord tones in music (Koelsch & Mulder, 2002; Patel, Gibson, Ratner, Besson, & Holcomb, 1998). Learners with German L1 might perceive the low accent 1 tone in the test sentences as out of key as it does not occur in a pragmatic context where it would be possible in German (i.e. polite, soothing requests (Grice, Baumann, & Benzmüller, 2005)). At the same time, the low Swedish tone is perceived frequently by the German speakers, but at this early stage in the learning process, there is no evidence that the connection between the tone and the inflectional morphology has been made. The rightward-going negativity within the slightly heterogeneous beginner group suggests that already in advanced beginners, increased familiarisation with a tone language brings about greater neural activations for pitch differences. Although the finding is based on a correlation analysis only and sub-group differences could not be found (presumably due to an insufficient number of participants for that purpose), it appears to suggest an early sensitisation to pitch patterns which is likely to be another important step in the acquisition of word accents.

4.3.1.2. Benefits and consequences of predictive processing: suffix effects. As regards processing benefits and consequences upon encountering an expected or unexpected suffix, the beginner learners in the present study produced none of the typical prediction-related ERP effects, i.e. neither a decreased N400 for successfully pre-activated suffixes nor an increased late positivity for unexpected, non-pre-activated suffixes. Together with the lack of a pre-activation negativity, these results indicate rather clearly that the beginner learner group has not yet reached the final, necessary stage in the acquisition of word accents: the word accent-morphology connection. This prevents them from being able to pre-activate suffixes at early stages in the learning process.

4.4. Implications for research on second language acquisition and on prediction in L2

The behavioural and neurophysiological data in this study suggest that complex predictive features like Swedish word accents are acquired in several consecutive stages, at least by learners from a non-tonal background. Acquisition of the predictive prosodic features does therefore not occur rapidly and initially, prediction cannot be used as a tool to facilitate language acquisition and processing.

Despite being a very frequent feature and non-existent in the learners' L1, it appears as though the Swedish stem tone-suffix system is too complex for beginner learners to acquire without first going through important intermediary stages. Discerning Swedish word accents and disentangling the association between these prosodic cues and the language's morphology is a lengthy step by step process which in turn renders the untaught prosody-morphology connection inaccessible for beginner learners. Notably, this feature of Swedish is not explicitly taught to either native speakers or L2 learners, implying its fully implicit acquisition through repetitive exposure.

Previous research has shown that simple meaning and phonotactic rules can be extracted quite rapidly in implicit second language learning (Gullberg et al., 2012; Gullberg et al., 2010; McLaughlin et al., 2004), which led us to believe that it might be possible for beginner L2 learners to master the prosodic cues to morphosyntactic information that Swedish word accents offer. Yet, it has also been shown that the implicit acquisition of second language items, i.e. words with non-native phonological features, is less rapid and automatic than the acquisition of novel L1 words (Kimppa et al., 2015, 2016). We thus propose that any lag in the acquisition of non-native linguistic items is based on the necessity of similar gradual dissociation, sensisation and (re-)association processes like those suggested for the current results. Complex non-native linguistic items appear to pose a great acquisition challenge for beginner learners resulting in the need for consecutive acquisitional stages.

Evidence of two such successive stages could be seen in this study. The presence of a neutral, pitch-related ERP response to the stem tones indicated that beginner L2s had learned to evaluate word accents as non-linguistic features rather than associating them with their L1 default pragmatic interpretation of pitch variation. This dissociation of tonal patterns from L1 pragmatic function arguably frees the tones for a re-association with L2 linguistic information. An extension of the pitch-related response which correlated positively with increasing language proficiency may be evidence that the L2 group in the present study is at the verge of arriving at a new important stage: the sensitisation to tonal differences in a new linguistic context. In order to associate a new linguistic function to the word tones, the learners need to learn to pay more attention to them, as well as focus on discerning different pitch patterns. Whether other additional stages are necessary before learners then associate word tones with their predictive morphological function is a question that the present study cannot answer as the beginner learner group never goes beyond the two earliest stages. The fact, however, that the pre-activating properties seem to be used already by upper intermediate learners (cf. behavioural findings in Schremm et al., 2016) suggests that the remaining acquisitional stages can progress rather rapidly. A learning study—either cross-sectional or longitudinal—with learners at different stages in the acquisition process as well as non-learners or tests at a pre-learning stage, respectively, could shed light on the entire process and test the hypotheses we have arrived at on the basis of the current data. It might similarly be informative to test speakers from a language that has no lexically or pragmatically meaningful pitch contours similar to those of accent 1 and accent 2. Although no lexical or pragmatic interference could be expected here, we assume that an intuitive search for linguistic relevance in the word tones might still take place. Learners could for instance interpret them as stress markers given that high tones are common as correlates of stress (e.g. Hayes, 1995; Lehiste, 1970), making some degree of dissociation and re-association necessary.

Regarding second language learners' capacity to predict L2 input, the present study's results indicate that it takes a relatively high proficiency before learners can reliably predict upcoming information based on available context. This is the case even with a feature that is seemingly optimal for second language learner prediction as it is highly frequent and invulnerable to L1 information processing strategies. We propose that this is due to a relatively slow gradual acquisition process resulting from a rather complex intertwining of morphosyntax, prosody and semantics in this predictive situation. In less complex contexts, L2 prediction may still be possible even for beginner learners. It could be rewarding to carry out a similar study in a less intricate predictive system. A possible candidate might be Dutch resyllabification in plural nouns ending in *-en* which changes both vowel duration and pitch pattern on the word stem (Kemps, Ernestus, & Schreuder, 2005). This would make prediction and consequently pre-activation of the upcoming suffix or lack thereof possible already at the word stem much like the Swedish word accent-suffix predictive pattern. Yet the system appears less complex, as it is delimited to certain well-defined morphosyntactic contexts. In such a situation, a much more rapid progression through the different acquisitional stages or a lack of those stages altogether are realistic possibilities.

5. Conclusion

In the present study, we found that beginner learners of Swedish, although not yet able to use the pre-activating properties of Swedish word accents, have mastered several precursory steps possibly leading up to the eventual acquisition of the pre-activation strategy. The learners showed no behavioural differences in response to either the stem tones or their validity with respect to following suffixes, and unlike native speakers, they did not produce any of the prototypical, word accent-suffix associated ERP effects: PrAN, N400 and P600. Instead, the L2 group's data yielded a mid-distributed negativity 400–600 ms after word accent onset which was virtually identical to a negativity they produced for pure pitch differences and which expanded right-laterally with increasing proficiency. These findings suggest that beginner learners of Swedish have learned to dissociate the word tonal patterns from the pragmatic function in their L1 and become increasingly sensitive to the pitch differences in Swedish. Both the dissociation and sensitisation can be seen as important and necessary steps before the tones can be associated with suffixes and attain their predictive capacity. Thus, beginner learners of a language are not able to immediately make use of complex predictive processing cues, when those are embedded in a complex interplay of prosody, semantics and morphosyntax. This holds true even if the cues are very frequent and unchallenged by direct counterparts in the learners' L1. The necessary precursory stages in the acquisition of complex non-native prosody-morphosyntax connections prevent them from being acquired as rapidly as has been shown for relatively more straightforward phonotactic patterns and sound-meaning correspondences. Related behavioural results suggest, however, that L2 learners at slightly more advanced levels of proficiency can master all necessary precursory stages and thus have access to the full predictive potential of such prosodic cues.

Acknowledgements

The work was supported by the Swedish Research Council (grant number 2011–2284), Knut and Alice Wallenberg Foundation (grant number 2014.0139), Lundbeck Foundation (Denmark; project 15480 Neolex; R140-2013-12951), RF Government (grant contract N° 14.W03.31.0010), and Marcus and Amalia Wallenberg Foundation (grant number 2014.0039). The authors gratefully acknowledge Lund University Humanities Lab.

Appendix A. Overview of word accent stimuli

NP	Verb ('got')	Target		PP	Validity
		Stem + Tone	Suffix		
Rut	fick	fisk$_{1/2}$	-en/-ar	till lunch	valid
Rut	fick	fisk$_{2/1}$	-en/-ar	till lunch	invalid
Ruth	*got*	*fish*	*-DEF.SG/PL*	*for lunch*	
Brit	fick	bock$_{1/2}$	-en/-ar	till jul	valid
Brit	fick	bock$_{2/1}$	-en/-ar	till jul	invalid
Bridget	*got*	*ram*	*-DEF.SG/PL*	*for Christmas*	
Kurt	fick	bult$_{1/2}$	-en/-ar	på da'n	valid
Kurt	fick	bult$_{2/1}$	-en/-ar	på da'n	invalid
Curt	*got*	*bolt*	*-DEF.SG/PL*	*on the day*	
Bengt	fick	bunt$_{1/2}$	-en/-ar	på sta'n	valid
Bengt	fick	bunt$_{2/1}$	-en/-ar	på sta'n	invalid
Ben	*got*	*bundle*	*-DEF.SG/PL*	*in town*	
Knut	fick	båt$_{1/2}$	-en/-ar	till jul	valid
Knut	fick	båt$_{2/1}$	-en/-ar	till jul	invalid
Knut	*got*	*boat*	*-DEF.SG/PL*	*for Christmas*	
Rut	fick	bänk$_{1/2}$	-en/-ar	på sta'n	valid
Rut	fick	bänk$_{2/1}$	-en/-ar	på sta'n	invalid
Ruth	*got*	*bench*	*-DEF.SG/PL*	*in town*	
Brit	fick	fläck$_{1/2}$	-en/-ar	på sta'n	valid
Brit	fick	fläck$_{2/1}$	-en/-ar	på sta'n	invalid
Bridget	*got*	*stain*	*-DEF.SG/PL*	*in town*	
Kurt	fick	flock$_{1/2}$	-en/-ar	på skoj	valid
Kurt	fick	flock$_{2/1}$	-en/-ar	på skoj	invalid

(*continued on next page*)

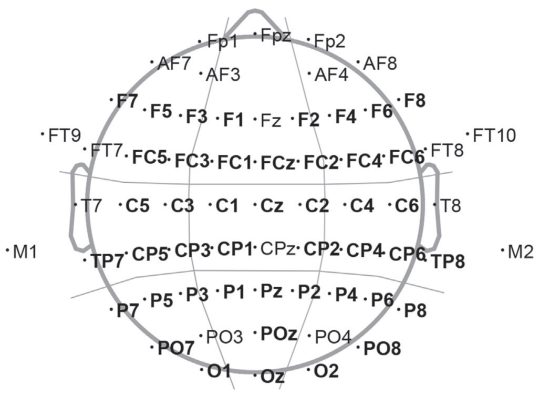
(*continued*)

NP	Verb ('got')	Target		PP	Validity
		Stem + Tone	Suffix		
Curt	*got*	*herd*	*-DEF.SG/PL*	*for fun*	
Bengt	fick	$\text{hatt}_{1/2}$	-en/-ar	till jul	valid
Bengt	fick	$\text{hatt}_{2/1}$	-en/-ar	till jul	invalid
Ben	*got*	*hat*	*-DEF.SG/PL*	*for Christmas*	
Knut	fick	$\text{hink}_{1/2}$	-en/-ar	på da'n	valid
Knut	fick	$\text{hink}_{2/1}$	-en/-ar	på da'n	invalid
Knut	*got*	*bucket*	*-DEF.SG/PL*	*on the day*	
Rut	fick	$\text{häst}_{1/2}$	-en/-ar	till jul	valid
Rut	fick	$\text{häst}_{2/1}$	-en/-ar	till jul	invalid
Ruth	*got*	*horse*	*-DEF.SG/PL*	*for Christmas*	
Brit	fick	$\text{klack}_{1/2}$	-en/-ar	till påsk	valid
Brit	fick	$\text{klack}_{2/1}$	-en/-ar	till påsk	invalid
Bridget	*got*	*heel*	*-DEF.SG/PL*	*for Easter*	
Kurt	fick	$\text{klick}_{1/2}$	-en/-ar	till lunch	valid
Kurt	fick	$\text{klick}_{2/1}$	-en/-ar	till lunch	invalid
Curt	*got*	*dollop*	*-DEF.SG/PL*	*for lunch*	
Bengt	fick	$\text{kock}_{1/2}$	-en/-ar	till påsk	valid
Bengt	fick	$\text{kock}_{2/1}$	-en/-ar	till påsk	invalid
Ben	*got*	*chef*	*-DEF.SG/PL*	*for Easter*	
Knut	fick	$\text{tupp}_{1/2}$	-en/-ar	till lunch	valid
Knut	fick	$\text{tupp}_{2/1}$	-en/-ar	till lunch	invalid
Knut	*got*	*rooster*	*-DEF.SG/PL*	*for lunch*	
Rut	fick	$\text{kork}_{1/2}$	-en/-ar	på skoj	valid
Rut	fick	$\text{kork}_{2/1}$	-en/-ar	på skoj	invalid
Ruth	*got*	*cork*	*-DEF.SG/PL*	*for fun*	
Brit	fick	$\text{krock}_{1/2}$	-en/-ar	på da'n	valid
Brit	fick	$\text{krock}_{2/1}$	-en/-ar	på da'n	invalid
Bridget	*got*	*crash*	*-DEF.SG/PL*	*on the day*	
Kurt	fick	$\text{kåk}_{1/2}$	-en/-ar	till jul	valid
Kurt	fick	$\text{kåk}_{2/1}$	-en/-ar	till jul	invalid
Curt	*got*	*shack*	*-DEF.SG/PL*	*for Christmas*	
Bengt	fick	$\text{lek}_{1/2}$	-en/-ar	på da'n	valid
Bengt	fick	$\text{lek}_{2/1}$	-en/-ar	på da'n	invalid
Ben	*got*	*play*	*-DEF.SG/PL*	*on the day*	
Knut	fick	$\text{låt}_{1/2}$	-en/-ar	på skoj	valid
Knut	fick	$\text{låt}_{2/1}$	-en/-ar	på skoj	invalid
Knut	*got*	*song*	*-DEF.SG/PL*	*for fun*	
Rut	fick	$\text{mink}_{1/2}$	-en/-ar	på da'n	valid
Rut	fick	$\text{mink}_{2/1}$	-en/-ar	på da'n	invalid
Ruth	*got*	*mink*	*-DEF.SG/PL*	*on the day*	
Brit	fick	$\text{lök}_{1/2}$	-en/-ar	till lunch	valid
Brit	fick	$\text{lök}_{2/1}$	-en/-ar	till lunch	invalid
Bridget	*got*	*onion*	*-DEF.SG/PL*	*for lunch*	
Kurt	fick	$\text{mack}_{1/2}$	-en/-ar	på sta'n	valid
Kurt	fick	$\text{mack}_{2/1}$	-en/-ar	på sta'n	invalid
Curt	*got*	*petrol station*	*-DEF.SG/PL*	*in town*	
Bengt	fick	$\text{ost}_{1/2}$	-en/-ar	till lunch	valid
Bengt	fick	$\text{ost}_{2/1}$	-en/-ar	till lunch	invalid
Ben	*got*	*cheese*	*-DEF.SG/PL*	*for lunch*	

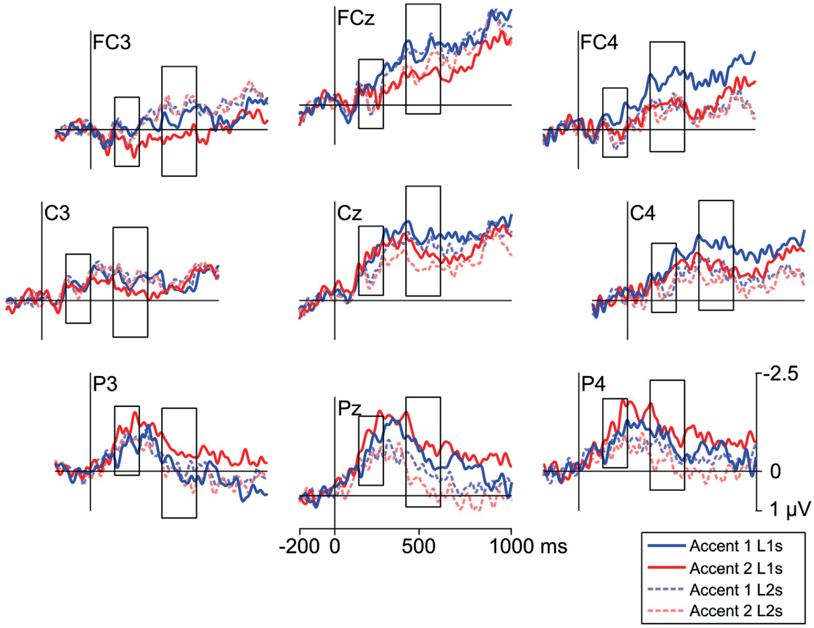
(*continued*)

NP	Verb ('got')	Target		PP	Validity
		Stem + Tone	Suffix		
Knut	fick	$prick_{1/2}$	-en/-ar	på sta'n	valid
Knut	fick	$prick_{2/1}$	-en/-ar	på sta'n	invalid
Knut	*got*	*dot*	*-DEF.SG/PL*	*in town*	
Rut	fick	$säck_{1/2}$	-en/-ar	till påsk	valid
Rut	fick	$säck_{2/1}$	-en/-ar	till påsk	invalid
Ruth	*got*	*bag*	*-DEF.SG/PL*	*for Easter*	
Brit	fick	$tolk_{1/2}$	-en/-ar	på skoj	valid
Brit	fick	$tolk_{2/1}$	-en/-ar	på skoj	invalid
Bridget	*got*	*interpreter*	*-DEF.SG/PL*	*for fun*	
Kurt	fick	$kopp_{1/2}$	-en/-ar	till påsk	valid
Kurt	fick	$kopp_{2/1}$	-en/-ar	till påsk	invalid
Curt	*got*	*cup*	*-DEF.SG/PL*	*for Easter*	
Bengt	fick	$bäck_{1/2}$	-en/-ar	på skoj	valid
Bengt	fick	$bäck_{2/1}$	-en/-ar	på skoj	invalid
Ben	*got*	*brook*	*-DEF.SG/PL*	*for fun*	
Knut	fick	$spark_{1/2}$	-en/-ar	till påsk	valid
Knut	fick	$spark_{2/1}$	-en/-ar	till påsk	invalid
Knut	*got*	*kick*	*-DEF.SG/PL*	*for Easter*	

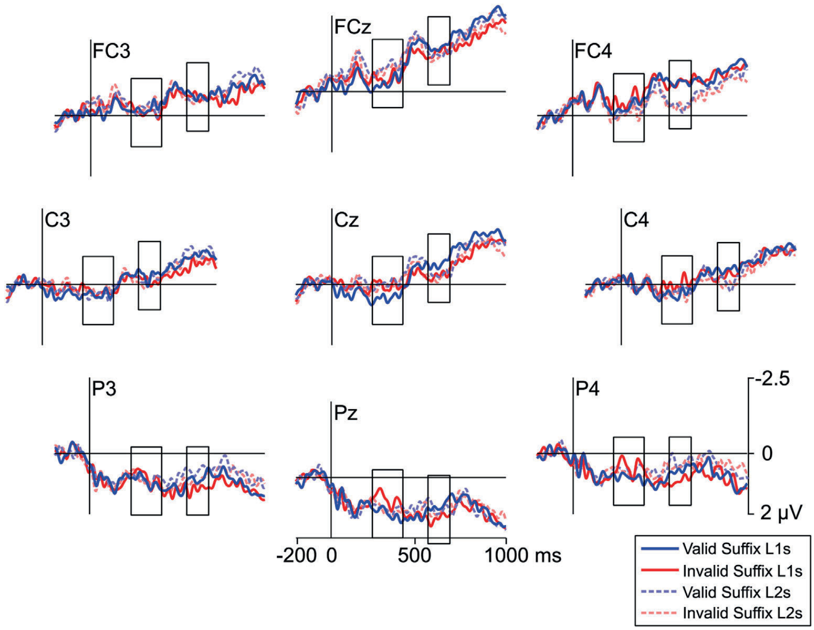
Appendix B. Complete overview of channel locations



Appendix C. ERP activity after stem tone onset at one electrode from each topographical region



Appendix D. ERP activity after suffix onset at one electrode from each topographical region



References

Bruce, G. (1977). *Swedish word accents in sentence perspective*. Lund: Gleerups.

DeJong, G. (1979). Prediction and substantiation: A new approach to natural language processing. *Cognitive Science, 3*, 251–273. https://doi.org/10.1207/s15516709cog0303_4.

DeLong, K. A., Urbach, T. P., & Kutas, M. (2005). Probabilistic word pre-activation during language comprehension inferred from electrical brain activity. *Nature Neuroscience, 8*(8), 1117–1121. https://doi.org/10.1038/nn1504.

Federmeier, K. D., Wlotko, E. W., De Ochoa-Dewald, E., & Kutas, M. (2007). Multiple effects of sentential constraint on word processing. *Brain Research, 1146*, 75–84. https://doi.org/10.1016/j.brainres.2006.06.101.

Fischler, I., & Bloom, P. A. (1979). Automatic and attentional processes in the effects of sentence contexts on word recognition. *Journal of Verbal Learning and Verbal Behavior, 18*, 1–20. https://doi.org/10.1016/S0022-5371(79)90534-6.

Foucart, A., Martin, C. D., Moreno, E. M., & Costa, A. (2014). Can bilinguals see it coming? Word anticipation in L2 sentence reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 40*(5), 1461–1469. https://doi.org/10.1037/a0036756.

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

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

Grice, M., Baumann, S., & Benzmüller, R. (2005). German intonation in autosegmental-metrical phonology. In S.-A. Jun (Ed.), *Prosodic typology: The phonology of intonation and phrasing* (pp. 55–83). Oxford: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199249633.003.0003.

Guillelmon, D., & Grosjean, F. (2001). The gender marking effect in spoken word recognition: The case of bilinguals. *Memory & Cognition, 29*(3), 503–511. https://doi.org/10.3758/BF03196401.

Gullberg, M., Roberts, L., & Dimroth, C. (2012). What word-level knowledge can adult learners acquire after minimal exposure to a new language?. In P. Jordens, & L. Roberts (Eds.), *International review of applied linguistics in language teaching* (Vol. 50, pp. 239–276). https://doi.org/10.1515/iral-2012-0010.

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

Hahne, A., & Friederici, A. D. (1999). Electrophysiological evidence for two steps in syntactic analysis: Early automatic and late controlled processes. *Journal of Cognitive Neuroscience, 11*(2), 194–205. https://doi.org/10.1162/089892999563328.

Hasegawa, M., Carpenter, P. A., & Just, M. A. (2002). An fMRI study of bilingual sentence comprehension and workload. *NeuroImage, 15*, 647–660. https://doi.org/10.1006/nimg.2001.1001.

Hayes, B. (1995). *Metrical stress theory: Principles and case studies*. Chicago: University of Chicago Press.

Hirst, D., & Di Cristo, A. (1998). A survey of intonation systems. In D. Hirst, & A. Di Cristo (Eds.), *Intonation systems: A survey of twenty languages* (pp. 1–44). Cambridge: Cambridge University Press.

Hollingshead, A. B. (1975). *Four factor index of social status*. Unpublished manuscript. New Haven, CT: Yale University.

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.

Jackendoff, R. (2002). *Foundations of language: Brain, meaning, grammar, evolution*. Oxford: Oxford University Press.

Jung, T.-P., Makeig, S., Humphries, C., Lee, T.-W., McKeown, M., Iragui, V., et al. (2000). Removing electroencephalographic artifacts by blind source separation. *Psychophysiology, 37*, 136–178. https://doi.org/10.1111/1469-8986.3720163.

Kaan, E. (2014). Prective 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. *Journal of Cognitive Neuroscience, 15*(1), 98–110. https://doi.org/10.1162/089892903321107855.

Kemps, R. J., Ernestus, M., & Schreuder, R. B. (2005). Prosodic cues for morphological complexity: The case of Dutch plural nouns. *Memory & Cognition, 33*(3), 430–460. https://doi.org/10.3758/BF03193061.

Kimppa, L., Kujala, T., Leminen, A., Vainio, M., & Shtyrov, Y. (2015). Rapid and automatic speech-pecific learning mechanisms in human neocortex. *NeuroImage*, 282–291. https://doi.org/10.1016/j.neuroimage.2015.05.098.

Kimppa, L., Kujala, T., & Shtyrov, Y. (2016). Individual language experience modualtes rapid formation of cortical memory circuits for novel words. *Scientific Reports, 6*(30227), 1–10. https://doi.org/10.1038/srep30227.

Koelsch, S., & Mulder, J. (2002). Electric brain responses to inappropriate harmonies during listening to expressive music. *Clinical Neurophysiology, 113*, 862–869. https://doi.org/10.1016/S1388-2457(02)00050-0.

Koester, D., Gunter, T. C., Wagner, S., & Friederici, A. D. (2004). Morphosyntax, prosody, and linking elements: The auditory processing of German nominal compounds. *Journal of Cognitive Neuroscience, 16*(9), 1647–1668. https://doi.org/10.1162/0898929042568541.

Kuperberg, G. R. (2013). The proactive comprehender: What event-related potentials tell us about the dynamics of reading comprehension. In B. Miller, L. E. Cuttling, & P. McCardle (Eds.), *Unraveling the behavioral, neurobiological, and genetic components of reading comprehension* (pp. 2–7). Baltimore: Paul Brookes Publishing.

Kuperberg, G. R., & Jaeger, F. (2016). What do we mean by prediction in language comprehension. *Language, Cognition and Neuroscience, 31*(1), 32–59. https://doi.org/10.1080/23273798.2015.1102299.

Kutas, M., DeLong, K. A., & Smith, N. J. (2011). A look around at what lies ahead: Prediction and predictability in language processing. In M. Bar (Ed.), *Predictions in the brain: Using our past to generate a future* (pp. 190–207). New York: Oxford University Press.

Lehiste, I. (1970). *Suprasegmentals*. Cambridge: MIT Press.

Leminen, A., Kimppa, L. L., 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.

Martin, C. D., Thierry, G., Kuipers, J.-R., Boutonnet, B., Foucart, A., & Costa, A. (2013). Bilinguals reading in their second language do not predict upcoming words as native readers do. *Journal of Memory and Language, 69*, 574–588. https://doi.org/10.1016/j.jml.2013.08.001.

McLaughlin, J., Osterhout, L., & Kim, A. (2004). Neural correlates of second-language word learning: Minimal instruction produces rapid change. *Nature Neuroscience, 7*(7), 703–704. https://doi.org/10.1038/nn1264.

Mestres-Missé, A., Rodriguez-Fornells, A., & Münte, T. F. (2007). Watching the brain during meaning acquisition. *Cerebral Cortex*, 1858–1866. https://doi.org/10.1093/cercor/bhl094.

Molinaro, N., Barber, H. A., Caffarra, S., & Carreiras, M. (2014). On the left anterior negativity (LAN): The case of morphosyntactic agreement. *Cortex*, 1–4. https://doi.org/10.1016/j.cortex.2014.06.009.

Norris, D. (1994). Shortlist: A connectionist model of continuous speech recognition. *Cognitition, 52*, 189–234. https://doi.org/10.1016/0010-0277(94)90043-4.

Patel, A. D., Gibson, E., Ratner, J., Besson, M., & Holcomb, P. J. (1998). Processing syntactic relations in language and music: An event-related potentital study. *Journal of Cognitive Neuroscience, 10*(6), 717–733. https://doi.org/10.1162/089892998563121.

van Pretten, C., & Luka, B. J. (2012). Prediction during language comprehension: Benefits, costs, and ERP components. *International Journal of Psychophysiology*, 176–190. https://doi.org/10.1016/j.ijpsycho.2011.09.015.

Riad, T. (1998). The origin of Scandinavian tone accents. *Diachronica, 15*(1), 63–98. https://doi.org/10.1075/dia.15.1.04ria.

Rischel, J. (1963). Morphemic tone and word tone in Eastern Norwegian. *Phonetica, 10*, 154–164. https://doi.org/10.1159/000258166.

Roll, M. (2015). A neurolinguistic study of South Swedish word accents: Electrical brain potentials in nouns and verbs. *Nordic Journal of Linguistics*, 149–182. https://doi.org/10.1017/S0332586515000189.

Roll, M., Gosselke, S., Lindgren, M., & Horne, M. (2013). Time-driven effects on processing grammatical agreement. *Frontiers in Psychology, 4*(1004), 1–8. https://doi.org/10.3389/fpsyg.2013.01004.

Roll, M., Horne, M., & Lindgren, M. (2010). Word accents and morphology - ERPs of Swedish word processing. *Brain Research, 1330*, 114–123. https://doi.org/10.1016/j.brainres.2010.03.020.

Roll, M., Söderström, P., & Horne, M. (2013). Word-stem tones cue suffixes in the brain. *Brain Research, 1520*, 116–120. https://doi.org/10.1016/j.brainres.2013.05.013.

Roll, M., Söderström, P., Mannfolk, P., Shtyrov, Y., Johansson, M., van Westen, D., et al. (2015). Word tones cueing morphosyntactic structure: Neuroanatomical substrates and activation time-course assessed by EEG and fMRI. *Brain & Language*, 14–21. https://doi.org/10.1016/j.bandl.2015.07.009.

Schremm, A., Söderström, P., Horne, M., & Roll, M. (2016). Implicit acquisition of tone-suffix connections in L2 learners of Swedish. *The Mental Lexicon, 11*(1), 55–75. https://doi.org/10.1075/ml.11.1.03sch.

Shtyrov, Y. (2011). Fast mapping of novel word forms traced neurophysiologically. *Frontiers in Psychology, 2*, 1–9. https://doi.org/10.3389/fpsyg.2011.00340.

Shtyrov, Y., Nikulin, V. V., & Pulvermüller, F. (2010). Rapid cortical plasticity underlying novel word learning. *The Journal of Neuroscience*, 16864–16867. https://doi.org/10.1523/JNEUROSCI.1376-10.2010.

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

Söderström, P., Horne, M., & Roll, M. (2016a). Word accents and phonological neighbourhood as predictive cues in spoken language comprehension. *Proceedings Speech Prosody, 2016*.

Söderström, P., Horne, M., & Roll, M. (2016b). Stem tones pre-activate suffixes in the brain. *Journal of Psycholinguistic Research*, 1–10. https://doi.org/10.1007/s10936-016-9434-2.

Söderström, P., Roll, M., & Horne, M. (2012). Processing morphologically conditioned word accents. *The Mental Lexicon, 7*(1), 77–89. https://doi.org/10.1075/ml.7.1.04soe.

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.

Wicha, N. Y., Moreno, E. M., & Kutas, M. (2003). Expecting gender: An event related brain potential study on the role of grammatical gender in comprehending a line drawing within a written sentence in Spanish. *Cortex, 39*, 483–508. https://doi.org/10.1016/S0010-9452(08)70260-0.

Williams, S. M. (2013). *A major revision of the Edinburgh handedness inventory*. https://doi.org/10.13140/2.1.5180.9284.

Paper II

Phonological transfer effects in novice learners: A learner's brain detects grammar errors only if the language sounds familiar

Abstract

Many aspects of a new language, including grammar rules, can be acquired and accessed within minutes. In the present study, we investigate how initial learners respond when the rules of a novel language are not adhered to. Through spoken word-picture association-learning, tonal and non-tonal speakers were taught artificial words. Along with lexicosemantic content expressed by consonants, the words contained grammatical properties embedded in vowels and tones. Pictures that were mismatched with any of the words' phonological cues elicited an N400 in tonal learners. Non-tonal learners only produced an N400 when the mismatch was based on a word's vowel or consonants, not the tone. The emergence of the N400 might indicate that error processing in L2 learners (unlike canonical processing) does not initially differentiate between grammar and semantics. Importantly, only errors based on familiar phonological cues evoked a mismatch-related response, highlighting the importance of phonological transfer in initial second language acquisition.

Introduction

Second language learners can acquire many aspects of a new language (L2) at surprisingly fast rates. In both naturalistic and artificial acquisition settings, novice learners show phonological (e.g., Gullberg, Roberts, Dimroth, Veroude & Indefrey, 2010), lexicosemantic (e.g., Dittinger, Scherer, Jäncke, Besson & Elmer, 2019; Gullberg et al., 2010; Kimppa, Kujala, Leminen, Vainio & Shtyrov, 2015), and even grammatical L2 knowledge (e.g., Cunillera et al., 2009; de Diego-Balaguer, Toro, Rodríguez-Fornells & Bachoud-Lévi, 2007; Gosselke Berthelsen, Horne, Shtyrov

& Roll, 2020) after mere minutes of language contact or training. An important facilitating factor during such early stages of second language acquisition (SLA) is transfer. To this effect, initial learners often acquire words with native language (L1) morphology and phonology more accurately and quickly than words with unfamiliar morphological and phonological features (Havas, Taylor, Vaquero, de-Diego-Balaguer, Rodríguez-Fornells & Davis, 2018; McKean, Letts & Howard, 2013). This is likely due to a more comprehensive neural processing of L1-like novel words which can draw on fine-tuned L1 networks. This claim finds support in electroencephalographic (EEG) studies which show that only words with familiar phonology induce rapid automatic word assessment and other pre-attentive processing mechanisms in initial learners (Gosselke Berthelsen et al., 2020; Kimppa et al., 2015). In the present paper, we extend the investigation of both initial L2 processing and the role of transfer by examining how novice learners initially process visually presented referents that mismatch lexicosemantic or morphosyntactic properties of acoustically presented words in a new language context. We investigate whether the learners respond differently to mismatches that are based on familiar and unfamiliar phonological cues.

Lexicosemantic and grammatical errors in L1 and L2

When a native speaker encounters a lexicosemantic or grammatical mismatch in the incoming language stream, their brain reacts to these errors with characteristic responses that can easily be observed using EEG measurements. A lexicosemantic error, for instance, will increase the amplitude of the N400 component. The N400 is a natural response to meaningful stimuli measured over the posterior part of the scalp. Its amplitude is in an inverse relation to the expectancy of an encountered stimulus with respect to context and world knowledge (Ganis, Kutas & Sereno, 1996; Kutas & Federmeier, 2014; Kutas & Hillyard, 1980). Hence, the component is particularly enhanced for lexicosemantic errors and can serve as an indicator of lexical and semantic accuracy. Sometimes, the N400 has even been observed for morphosyntactic manipulations (Barber & Carreiras, 2005; Guarjardo & Wicha, 2014). In those cases, the morphosyntactic irregularities are either embedded in contexts that also include prominent semantic manipulations (Guarjardo & Wicha, 2014) or constructions with a strong semantic load (i.e., N400 for noun-postnominal adjective agreement vs. left anterior negativity [LAN] for determiner-noun agreement, Barber & Carreiras, 2005). Most often, however, morphosyntactic processing is distinct from semantic processing. It typically evokes one or several of a range of sequential event-related potentials (ERPs): early left anterior negativity (ELAN), LAN, and/or P600. These components are enhanced and thus most easily elicited in contexts where specific parts of morphosyntactic processing are manipulated so as to fail, i.e., in error contexts. The suggested functions of the different morphosyntactic components are as follows: The ELAN, with a latency of ~150 ms, is believed to signal the assessment of a word's grammatical category, presumably based on the initial activation of its morphosyntactic features (e.g., Friederici, 2002; Gosselke Berthelsen et al., 2020; Neville, Nicol, Barss, Forster & Garret, 1991). The LAN, subsequently, is similar to the lexicosemantic N400 in timing, but not distribution. It is largest at frontal electrodes, often left-lateralised, and is likely reflective of agreement assessment of, for instance, gender or number relations (Molinaro, Barber, Caffarara & Carreiras, 2014; Roll, Gosselke, Lindgren & Horne, 2013; Schremm, Novén, Horne & Roll, 2019). This assessment is likely made possible by a rule-based integration of morphosyntactic and semantic content (Gosselke Berthelsen et al., 2020; Krott & Lebib, 2013). The P600, finally, is a positive response strongest at posterior electrodes at around 600 ms. It reflects utterance-level integration processes and is enhanced in contexts of encountered errors (morphosyntactic and sometimes semantic) due to a need of revision or repair (Kim & Osterhout, 2005; Osterhout & Holcomb, 1992; Roll, Horne & Lindgren, 2007). Given their different but often complementary assumed functions, the three major morphosyntactic components can occur together or separately. For the most part, they stand in clear contrast to the lexicosemantic N400.

Error processing has also been studied in second language learners. With respect to lexicosemantic mismatches in the L2, learners resemble native speakers and produce an N400 already at low proficiency levels. With respect to morphosyntactic processing, however, learners differ profoundly from native speakers. Rather than an ELAN, LAN, or P600 for grammatical irregularities, learners often produce an N400, especially at low proficiency levels (McLaughlin et al., 2010; Tanner, McLaughlin, Herschensohn & Osterhout, 2012). The N400 initially acts as a second language learner's default response to any linguistic inconsistency, (morpho)syntax included. During continued acquisition, language-specific regularities gradually become grammaticalised in the learner's brain (Steinhauer, 2014): at intermediate proficiencies, learners produce P600 responses (e.g., Tanner et al., 2012; Tokowicz & MacWhinney, 2005), followed by the LAN at high proficiency levels (e.g., Bowden, Steinhauer, Sanz & Ullman, 2013; Gillon Dowens, Vergara, Barber & Carreiras, 2010; Ojima, Nakata & Kakigi, 2005), and, very rarely, an ELAN (e.g., Hanna, Shtyrov, Williams & Pulvermüller, 2016; Hed, Schremm, Horne & Roll, 2019; Rossi, Gugler, Friederici & Hahne, 2006). Evidently, however, language processing can be accelerated dramatically in targeted experimental settings. Here, learners frequently show both lexicosemantic and morphosyntactic ERP effects with minimal instruction or exposure. The order in which the ERP components emerge stays the same. Yet, the P600 emerges already after hours or even minutes of training in experimental settings (Batterink & Neville, 2013; Davidson & Indefrey, 2009; de Diego-Balaguer et al., 2007; Friederici, Steinhauer & Pfeifer, 2002; Grey, Sanz, Morgan-Short & Ullman, 2018; Havas, Laine & Rodriguez-Fornells, 2017; Mueller, Hahne, Fujii & Friederici, 2005). The same holds true for the LAN and possibly the ELAN which, however, seem to depend more on intermitted consolidation periods (Friederici, Steinhauer & Pfeifer, 2002; Mueller, Hirotani & Friederici, 2007; Morgan-Short, Steinhauer, Sanz & Ullman, 2012). While a recent artificial learning study found an ELAN and a LAN without consolidation, this may well be explained by the fact that this study analysed canonical grammar processing rather than error processing (Gosselke Berthelsen et al., 2020). Collectively, the findings for error processing in L2 learners suggest that lexicosemantic error processing appears to precede grammatical processing and that the latter proceeds gradually. Learners can process L2 words like native speakers but it either takes time or specific, experimental focus for the error-related ERP responses to emerge. In experimental settings, language processing can emerge rapidly, even in initial learners and often, in fact, before behavioural performance transcends chance levels (cf., McLaughlin, Osterhout & Kim, 2004; Hed et al., 2019). However, some grammar-related ERP components (ELAN/LAN) have thus far only been observed after consolidation periods, at least in violation paradigms.

Transfer in the early stages of SLA and word-level tone

One factor that appears critical for how rapidly and naturally novel items and rules can be acquired and processed is transfer from the learner's native language on the basis of L1-L2 similarity. Transfer is relevant on many linguistic levels, including syntax (e.g., Foucart & Frenck-Mestre, 2011; 2012) and phonology (e.g., Gosselke Berthelsen et al., 2020; Kimppa et al., 2015). Transfer tends to affect early, pre-conscious processing more than late, conscious processing or offline responses (e.g., Andersson, Sayehli & Gullberg 2019; Gillon Dowens et al., 2010; Gillon Dowens, Guo, Guo, Barber & Carreiras, 2011). With respect to phonological similarity, only novel words that are in accordance with the learners' L1 rules for sounds and sound structure are assessed and processed pre-attentively by initial learners. This has been found for a number of different pre-attentive responses: An early lexicosemantic gating component at ~50 ms, for instance, shows rapid lexical trace formation only for words that have L1-like phonotactic (Kimppa et al., 2015) or prosodic structure (Gosselke Berthelsen et al., 2020). For words where L1 transfer is not possible, the lexicosemantic gating component is unaffected. Similarly, the fronto-central N1, strongly involved in automatic auditory processing, is significantly reduced for phonologically illegal pseudowords but not for legal pseudowords (Silva, Vigário, Fernandez, Jerónimo, Alter & Frota, 2019). This may suggest that novel words with illegal phonology are not processed as words, but as unusual sound patterns. Finally, even pre-conscious grammar processing is affected by the possibility for phonological transfer. In a study on artificial word acquisition, Gosselke Berthelsen et al. (2020) found an ELAN response for pre-attentive morphosyntax activation only in learners for whom L1-L2 similarity was high. Together, these studies thus indicate that, at least initially, only words with L1-like phonology can be processed outside of consciousness. This is likely a result of listeners being able to extend their neural subsystems, fine-tuned to the automatic processing of native sounds, to native-sounding foreign words, in line with the concept of L1-L2 transfer. An exception to the illustrated limitations in pre-attentive processing has been observed in children (Partanen et al., 2017). Children produced changes in early components for non-native as well as native-sounding words. This is likely due to larger neural plasticity in the developing brain. However, the children still displayed clear hemispheric differences in the processing of phonologically native and non-native novel words, thus indicating that phonological similarity and transfer have an important impact on pre-attentive processing in all learners, even children. Presumably as a consequence of transfer and the possibility for pre-attentive processing, unfamiliar phonology also sometimes negatively affects behavioural learning outcomes (Havas et al., 2018; McKean et al., 2013) and impedes predictive processing (Lozano-Argüelles, Sagarra & Casillas, 2019; Sagarra & Casillas, 2018).

A grammar feature that lends itself well to the study of transfer in initial L2 acquisition is the so-called grammatical tone. Grammatical tone is a suprasegmental feature that is added onto the segments (vowels, consonants) of a word to make grammatical distinctions (e.g., case, gender, or number). Since the tone itself as well as its grammatical content are independent, additive features, the word's lexical and non-tonal grammatical meaning (e.g., suffixes) can be accessed irrespective of the tone. The tone's grammatical meaning, on the other hand, can only be accessed through the tone itself. If words with grammatical tone are taught to (beginner) learners from tonal and non-tonal language backgrounds, it is possible to directly compare general acquisition abilities (segmental features) to potentially transfer-affected acquisition (suprasegmental features) within the same words[1]. Importantly for the purpose of the current study, this can be extended to a direct comparison of initial learners' error processing (i.e., errors across segmental features) and error processing in light of potential transfer (errors in the suprasegmental tone dimension). An important premise, of course, is that L1-L2 transfer is possible for tone. This has, indeed, been observed behaviourally, at least when the L2 tone is less complex than the L1 tone and similar in form and function (e.g., Braun & Johnson, 2011; Hallé, Chang & Best, 2004; So & Best, 2010; Wayland & Li, 2008). Neurophysiological accounts of transfer effects in L2 learners of tone languages, on the other hand, are rare. However, studies using mismatch negativity (MMN) paradigms have observed transfer based modulations of L2 tone perception. To this effect, only learners with a tonal L1 showed modulations of MMN effects based on tone functionality and not only pitch height (Shen & Froud, 2019; Yu, et al., 2019). For post-perceptive processing, no N400 for lexical tone mismatches or LAN/P600 for mismatches in grammar-related tone is found in natural L2 learners with a non-

[1] It is possible, of course, that the mere presence of tone, despite its additive nature, poses an acquisition-hindrance for non-tonal learners. This would result in impoverished performance and processing in all conditions and would become apparent in the analysis of the segmental features.

tonal L1 (Gosselke Berthelsen et al., 2018; Pelzl, Lau, Guo, & DeKeyser, 2019). In addition to these null results, the clearest indication of transfer was found for artificial, grammatical tone in Gosselke Berthelsen et al. (2020): A rapidly emerging left anterior negativity (LAN), a P600, as well as early, pre-attentive processing (i.e., word recognition effect and ELAN) emerged in learners with a tonal L1. Non-tonal learners, in contrast, showed no evidence of pre-attentive processing and a later LAN onset. Thus, we see distinct transfer effects for L2 tone acquisition and processing. We, therefore, suggest that grammatical tone is well suited for examining how learners, differentially affected by transfer, differ in their processing of L2 errors.

Word-picture association learning and picture processing

While transfer is an important and often deliberately studied feature in initial L2 acquisition, studies attempt to limit the EXPLICIT presence of the learners' L1 in initial learning contexts to avoid a conscious mediation of the L2 through the L1. A useful method for circumventing overt L1 exposure during initial L2 learning is word-picture association learning where new L2 words are taught with the help of pictures. Interestingly, studies using this paradigm have uniformly found that picture associations can install rapid lexicosemantic (Dittinger et al., 2019; François, Cunillera, Garcia, Laine & Rodriguez-Fornells, 2017; Havas et al., 2017; Yang & Li, 2019) or grammatical (Gosselke Berthelsen et al., 2020; Havas et al., 2017) processing in novel words. For this to occur, learners must be aware of the lexical and grammatical content of the meaning-assigning pictures. Relatively little, however, is known about how this affects the processing of grammatically relevant features (e.g., number or gender) in meaning-carrying image referents themselves. We know from early N400 literature, that at least the N400 appears to be largely amodal and increases for inconsistent sentence endings regardless of whether they are expressed through words or pictures (e.g., Ganis et al., 1996; Kutas & van Petten, 1990; Nigam, Hoffman & Simons, 1992). In fact, even pictures primed with single auditory words show language-like N400 modulations (Pratarelli, 1994). Processing of grammatical properties in pictures, on the other hand, has been studied considerably less. One study that did investigate this found a LAN (but no P600) for gender mismatched pictures in sentential contexts (Wicha, Moreno & Kutas, 2003). This suggests that at least some lexicosemantic and morphosyntactic processes can be activated by pictures. Based on these findings, we assume that lexicosemantic and grammatical errors in the picture of a word-picture association paradigm should evoke the corresponding ERP responses (e.g., N400, LAN) and therefore constitute a useful tool for testing error processing in learners at the very beginning of the acquisition process.

The current study

Addressing a thus far understudied issue in initial second language acquisition, we investigated how error processing is realised during the initial acquisition of L2 words. As mentioned, learners at first process all L2 errors lexicosemantically. Grammatical components for error processing emerge later in both natural and artificial SLA. In artificial SLA, this process can, however, be sped up and grammatical processing can be found with minimal exposure. In the present study, we drew on a word-picture association paradigm that had previously produced rapid grammatical processing in novel words (Gosselke Berthelsen et al., 2020). The words in the present study contained grammatical tone and were tested on learners with and without a tonal L1. This allowed for a fine-grained study of phonological transfer effects. As a control condition, the words also contained a grammatical vowel change which both learner groups were equally familiar with from their L1. Importantly, we included mismatches in the word-picture pairs that became transparent at picture onset and analysed the learners' behavioural and neural responses to the resulting errors. Previous studies have illustrated that pictures in linguistic contexts can evoke both lexicosemantic and grammatical responses. Moreover, we were aware that word-picture association paradigms can successfully be used to teach grammar; a fact which entails that the linguistic content of pictures in this paradigm must be evident to learners. Therefore, pictures were considered a suitable, straightforward way of studying the processing of morphosyntactic mismatches in novice learners. Based on previous results and due to the strong focus on grammar in this paradigm, we expected grammatical responses (e.g., LAN/P600) to errors in the initial learners. We further anticipated error processing to be facilitated by L1-L2 transfer. To test for this, we elicited errors in a phonological dimension that one learner group was unfamiliar with (i.e., tone) or in dimensions that both learner groups knew from their L1 (i.e., vowel and consonant). We hypothesised that transfer effects in the form of reduced or missing mismatch responses would only appear in the unfamiliar tone condition, since the tone was an additive feature and the remainder of the word could be acquired independently of the tone.

Methods

Participants

For the present study, we recruited two groups of right-handed participants (assessed by revised Edinburgh Handedness Inventory; Williams, 2013): 24

participants with a native language with grammatically relevant tone, Swedish[2] , henceforth 'tonal L1s' (TL1s), as well as 24 participants with a closely-related non-tonal native language, German, henceforth 'non-tonal L1s' (NTL1s). The groups were matched for gender (TL1 = 10 males, NTL1 = 11 males), age (M_{TL1} = 23.7 years, *SD* = 2.6; M_{NTL1} = 23.7 years, S*D* = 1.6), socioeconomic status (Hollingshead, 1975), working memory span (tested via an automated operation span task, cf. Unsworth, Heitz, Schrock & Engle, 2005), as well as the number of languages spoken at above-elementary proficiency (M_{TL1} = 2, *SD* = 1; M_{NTL1} = 2, *SD* = 1). Furthermore, all participants had normal hearing defined as pure-tone hearing thresholds ≤20 dB Hearing Level (ISO, 2004) and could correctly discern tested auditory variations in linguistic input (i.e., vowel length, M_{TL1} = 97.7%, *SD* = 3.0; M_{NTL1} = 96.5%, *SD* = 4.0) as well as non-linguistic input (i.e., pitch, M_{TL1} = 91.2%, *SD* = 9.0; M_{NTL1} = 93.8%, *SD* = 6.2).

All experiments were conducted in agreement with the ethical guidelines for experiments in the Declaration of Helsinki and carried out in the Lund University Humanities Lab. All NTL1 participants were exchange students at Lund University. The NTL1s had no extensive knowledge of Swedish (highest self-assessed level of proficiency was B1 [=intermediate] of the Common European Framework of Reference, mean proficiency was A1 [=beginner]). Despite having lived in Sweden for some months (*M* = 23 weeks, *SD* = 16) and having studied Swedish to some extent (*M* = 10 weeks, *SD* = 7), they only engaged with Swedish actively for on average 5 hours per week (i.e., studying, conversation, listening, *SD* = 7) and passively for 10 hours (i.e., Swedish spoken in the background, *SD* = 8). Only one of the NTL1 participants reported having heard of Swedish word accents and none were aware of their ties to grammar nor did they include them in their vocabulary acquisition routines. Additionally, we tested the NTL1s on their perception of Swedish word accents and both behavioural and neuropsychological data suggested that they processed them as linguistically irrelevant.

Keeping with participant sizes in previous studies with related paradigms (e.g., 23: de Diego-Balaguer, Rodríguez-Fornells & Bacoud-Lévi, 2015; 19: Havas, Laine & Rodríguez-Fornells, 2017; Leminen, Kimppa, Leminen, Lehtonen, Mäkelä & Shtyrov, 2016), we originally recorded 24 participants per group. As each participant had their own set of stimuli, multiples of eight allowed us to counterbalance the distribution of the four grammatical features over vowels and tones. Prior to analyses, we excluded one participant from each group, due to rudimentary knowledge of a tone language (Chinese) and experienced discomfort with the EEG equipment, respectively. Closely inspecting the behavioural data, we

[2] Swedish has tones (word accents) that are strongly associated with grammatical suffixes. A low tone (accent 1) on the stem of the word *fisk* ('fish'), for instance, can be followed by the definite singular suffix *-en*, (*fisk*$_1$*-en*, fish-the, 'the fish') but not by the plural suffix *-ar* (**fisk*$_1$*-ar*, fish-PL, 'fish'). The plural instead requires a high tone (accent 2) on the word stem: fisk2-ar.

subsequently found that some participants performed poorly on the acquisition task of the main study and remained at chance level for error detection throughout the first session of the experiment (see section 2.3 below). Closer inspection of the distribution of the accuracy data (Figure 1) indicated a discontinuity in accuracy between 55 and 60% that we interpreted as a natural cut-off point. We therefore excluded the four participants (2 NTL1s, 2 TL1s) whose accuracy during the first session stayed below 55%, resulting in final group sizes of 21 participants.

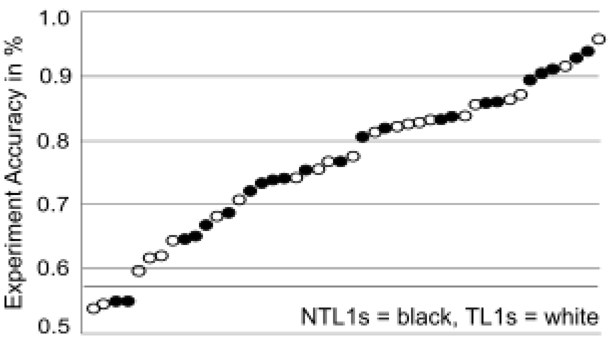


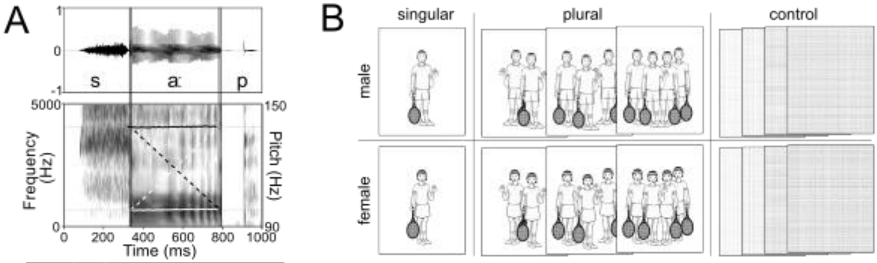
Figure 1. Distribution of participants' accuracy data during the first session, from low to high. Participant group indicated by black (TL1) and white (NTL1) dots.

Stimuli

We used a learning paradigm that has previously been reported to be efficient in learning novel lexical and morphosyntactic features with tonal and non-tonal distinctions (Gosselke Berthelsen et al., 2019). Participants learned new word-picture associations by hearing an auditorily presented novel word followed by a meaning-assigning picture, as described in more detail below:

Auditory stimuli

For the sound stimuli, we constructed pseudowords from consonants and vowels, which were produced by a native speaker of Russian (rather than German or Swedish to prevent differential carry-over effects) and chosen on the basis of being phonologically close in all three languages. Using Praat (Boersma, 2001), consonant and vowel durations were standardised. Subsequently, consonants and vowels were spliced into 24 simple CVC pseudowords and four different tones were added to the words with the help of pitch-manipulation. See Figure 2 for an overview of all words and an example stimulus.



		Vowels			
		a	ɛ	i	u
Initial consonants	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

Professions			
doctor	teacher	chemist	photographer
boxer	painter	race driver	basketball player
fisher	waiter	journalist	fire fighter
tailor	gardener	tennis player	flight attendant
cook	flautist	violinist	hairdresser
singer	cleaner	ballet dancer	police officer

Figure 2. A. Auditory word stimuli: Acoustic waveform, spectrogram and the four fundamental frequency (F0) curves for an example stimulus (above) and a list of all test words (below). B. Visual stimuli: An example set of picture stimuli for one profession (above) and a list of all professions in the experiment (below).

Picture stimuli

We chose images of humans portrayed as 24 different professions for the pictorial stimuli. For each profession, we constructed eight different images: one picture each with one, two, three, or four female or male workers, respectively. This way, in addition to the lexical information (profession), we were able to visually implement two grammatical categories: gender (masculine-feminine) and number (singular-plural). For the control condition, non-meaningful pictures were constructed by scrambling all pixels from the profession pictures. See Figure 2 for an overview of all professions and an example set of picture stimuli. See the supplementary material for an overview of all used picture stimuli.

Experimental procedure

All participants took part in two acquisition sessions, 24 hours apart. They sat on a chair one meter from a computer screen and kept their index fingers on a response box that stood on a table in front of them. 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. After a short instruction and training session (only in session 1), participants were asked to learn 24 words consisting of six professions that could be masculine or feminine (gender) and singular or plural (number). Professions were always expressed through the initial and final consonant (consonant frame). Gender and number were always expressed by vowel or tone, equally distributed across participants. Each participant had a unique set of words (i.e., list of six professions and combination of vowels, tones and their respective meaning). All vowels and tones were used equally often as part of target and control words. For an example set of stimuli, see Table 1.

Table 1. An example set of stimulus words.

	Target words								Control words (no meaning)		
	Masculine				**Feminine**						
	Singular		*Plural*		*Singular*		*Plural*				
violinist	d*i*f	**fall**	d*u*f	**fall**	d*i*f	**high**	d*u*f	**high**	dap	/ dɛp	rise/low
cleaner	f*i*f	**fall**	f*u*f	**fall**	f*i*f	**high**	f*u*f	**high**	fap	/ fɛp	rise/low
tennis player	k*i*p	**fall**	k*u*p	**fall**	k*i*p	**high**	k*u*p	**high**	kaf	/ kɛf	rise/low
painter	l*i*r	**fall**	l*u*r	**fall**	l*i*r	**high**	l*u*r	**high**	lap	/ lɛp	rise/low
gardener	s*i*s	**fall**	s*u*s	**fall**	s*i*s	**high**	s*u*s	**high**	sap	/ sɛp	rise/low
race driver	t*i*p	**fall**	t*u*p	**fall**	t*i*p	**high**	t*u*p	**high**	taf	/ tɛf	rise/low

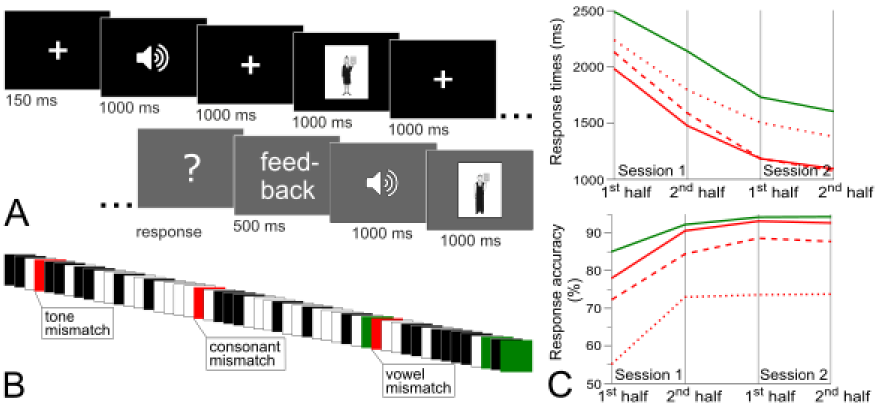


Figure 3. A. Experiment procedure: Black slides show test and control trials, grey slides are added for question trials. B. Example cycle from the experiment highlighting the random distribution of test trials (black), control trials (white), matched question trials (green), and mismatch trials (red). C. Behavioural results for the question trials: Matched question trials in green, mismatched trials in red. Consonant mismatch in solid lines, vowel mismatch in dashed lines, and tone mismatch in dotted lines.

Using the procedure shown in Figure 3, all 48 word-picture combinations from each participant's target and control set were presented 30 times in each of the two sessions. In every full repetition of all 48 experiment items (cycle), three pseudorandomly selected meaningful word-picture trials were followed by overt questions pertaining to the correctness of the previous word-picture pair. Every cycle also included three additional question trials where there was a mismatch between word and picture related to one of the learned categories: number, gender (i.e., based on tone or vowel) or profession (based on consonants). Participants responded to all question trials by pressing a button on the response box and overt feedback on their response was given. After every 10 cycles (~40 minutes), participants were offered a break.

Electroencephalography (EEG)

Throughout the experiment, 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 for the scalp channels were kept below 3 kΩ 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 for word and picture stimuli at word disambiguation point and picture onset, respectively. The extracted epochs included both learning and question trials. Independent component analysis (ICA) (Jung, Makeig, Humphries, Lee, McKeown, Iragui & Sejnowksi, 2000) was conducted on all epochs. ICA components representing eye artefacts and single bad channels were removed. Finally, all epochs still exceeding ±100 µV were excluded. Only the epochs pertaining to picture stimuli were considered for the analysis.

Statistical analysis

Behavioural data analysis was conducted separately for the two experimental factors 'Response Time' (RT) and 'Response Accuracy' which were recorded for question trials. RT measures were log-transformed to normalise the distribution and submitted to a mixed Analysis of Variance (ANOVA) in IBM SPSS Statistics 25 (International Business Machines Corp., Armonk, NY, United States) with within-subject factor 'Type' (consonant-related word-picture mismatch [henceforth, consonant mismatch], vowel-related word-picture mismatch [henceforth, vowel mismatch], and tone-related word-picture mismatch [henceforth, tone mismatch], and correctly matched word-picture trials [henceforth, matched pictures]), temporal factors 'Session' (session 1 vs. session 2) and 'Half' (first vs. second half of a session), and the between-subject factor 'Learner Group' (TL1s vs. NTL1s). For the response accuracy data, d′ scores were calculated for each participant, condition and time window by comparing z transforms of hit rates (correct acceptance of matched trials relative to the total number of matched trials) and false alarm rates (incorrect acceptance of mismatched trials compared to the total number of mismatched trials - by mismatch type). Log-linear corrections (Hautus, 1995) were applied to extreme values (i.e., 0 and 1). D′ scores were submitted to a mixed ANOVA with the factors 'Mismatch Type' (consonant, vowel, tone), the temporal factors 'Session' and 'Half' and the between-subject factor 'Learner Group'. Greenhouse-Geisser correction was used when applicable. Main effects and interactions were considered significant at a p-value of < 0.05. For pairwise comparisons, False Discovery Rate (FDR) corrections (Benjamini & Hochberg, 1995) were used.

To find effects of interest in the ERPs, global root mean squares (gRMS) of the data were used to investigate when the learner groups' neural activity was maximal (cf. Lehmann & Skrandies, 1980). When there were peak latency differences between the learner groups, we submitted the participants' unweighted mean gRMS peaks to a 2-tailed independent samples t-test. If there were significant between-group differences for gRMS peak latency, we established different time windows for the ERP analyses for the respective peak and analysed the groups separately. Subsequent to the gRMS analysis, we inspected the ERPs at gRMS peak latencies to find the ERP effects which the peaks were related to. The gRMS curve reaches a maximum when there is a CHANGE in neural activity and gRMS peaks therefore often mark effect onset rather than effect peak (cf. e.g., Roll, Söderström, Mannfolk, Shtyrov, Johansson, van Westen & Horne, 2015). Thus, using gRMS maxima as effect onsets, we selected time windows of 70 ms for the semi-early peaks (150 - 300 ms) and 200 ms for late peaks (above 300 ms) for the ERP analysis.

Testing for ERP effects at the gRMS peak latencies, mean amplitudes over the selected time windows for each electrode and condition were submitted to a cluster-based permutation test to find electrode groups that differed significantly between conditions. Each permutation test included one of the mismatched conditions and the match condition. We also tested for interactions with the between-subject factor 'Learner Group' as well as the temporal factors 'Session' and 'Half' and combinations thereof (e.g., interaction with 'Half' in session 1) in all conditions. All permutation analyses were carried out with help of the nonparametric cluster-based permutation approach implemented in Fieldtrip (Maris & Oostenveld, 2007). Using the Monte Carlo method to account for large data sets, we ran 1000 random permutations of the data. Clusters of three or more electrodes that had a p-value of < 0.05 were considered significant.

Finally, we conducted two-tailed Pearson correlations between response times and effect amplitudes (i.e., difference between mismatch and match amplitudes for electrodes in mismatch clusters) for all emerging error-related clusters in order to examine possible relationships between error processing in novice learners and subsequent behavioural responses.

Results

Behaviour

For the behavioural variable Response Accuracy, the mixed ANOVA produced a main effect of Mismatch Type, see Table 2. FDR corrected pairwise comparisons revealed that participants were significantly more accurate for vowel or consonant conditions than for tones. Further, mismatches in consonants were more accurately detected than those involving vowels. Learners improved throughout and between sessions, particularly from the beginning to the end of session 1. The non-tonal learners' tone detection ability was initially so poor that it differed clearly from vowel and consonant mismatches but increased over time and approached vowel mismatch levels towards the end of the learning sessions.

Table 2: Results of the mixed Analysis of Variance (ANOVA) analysis for Response Accuracy (d′ values) as well as means and standard variations of the raw data. (S1 = session 1, S2 = session 2, H1 = first half, H2 = second half)

Main effects, Interactions, Pairwise Comparisons			*F*	*p*	Means in % (*SD*)			
Mismatch Type			19.12	<.001				
Pairwise comparisons:	Tone	vs Vowel		<.001	Tone Mismatch			69.07 (26)
		vs Conson.		<.001	Vowel Mismatch			83.50 (14)
	Vowel	vs Conson.		.004	Cons. Mismatch			88.77 (9)
Session			96.65	<.001	S1	79.91 (9)	S2	87.46 (12)
Half			122.58	<.001	H1	80.19 (10)	H2	86.30 (11)
Session * Half			24.33	<.001				
Session 1: Half			145.44	<.001	H1	72.75 (11)	H2	85.28 (10)
Session 2: Half			43.61	<.001	H1	87.61 (12)	H2	87.31 (13)
Mismatch Type * Half * Learner Group			3.12	.050				
TL1s: Type * Half			0.34	.694				
NTL1s: Type * Half			3.51	.044				
Half 1: Mismatch Type			16.62	<.001				
Pairwise comp.: Tone		vs Vowel		<.001	Tone Mismatch			58.10 (25)
		vs Conson.		<.001	Vowel Mismatch			80.32 (19)
					Cons. Mismatch			84.44 (13)
Half 2: Mismatch Type			5.31	.017				
Pairwise comp.: Tone		vs Conson.		.045	Tone Mismatch			70.48 (31)
					Vowel Mismatch			82.54 (18)
					Cons. Mismatch			89.68 (10)

For Response Times, we found a main effect of Type (see Table 3). FDR corrected pairwise comparisons revealed that participants were faster at detecting consonant mismatches and vowel mismatches than tone mismatches or matched pictures (for means and p-values, refer to Table 3). Furthermore, Response Times improved throughout the sessions and between sessions. The strongest improvement took place within session 1. Vowel mismatch detection was not significantly faster than tone mismatch detection until session 2. There were no significant interactions with Learner Group.

Table 3: Results of the mixed Analysis of Variance (ANOVA) analysis for Response Times (log-transformations) as well as means and standard variations of the raw data. (T Mm = tone mismatch, V Mm = vowel mismatch, C Mm = consonant mismatch, S1 = session 1, S2 = session 2, H1 = first half, H2 = second half)

Main effects, Interactions, Pairwise Comparisons			*F*	*p*	Means in ms (*SD*)			
Type			39.71	<.001				
Pairwise comparisons:	T Mm	vs V Mm		<.001	Tone Mismatch			1731 (794)
		vs C Mm		<.001	Vowel Mismatch			1497 (705)
		vs Match		<.001	Cons. Mismatch			1437 (661)
	V Mm	vs Match		<.001	Match			1991 (701)
	C Mm	vs Match		<.001				
Session			49.50	<.001	S1	1982 (773)	S2	1346 (713)
Half			39.07	<.001	H1	1807 (670)	H2	1521 (720)
Session * Half			11.83	<.001				
Session 1: Half			34.69	<.001	H1	2214 (807)	H2	1751 (865)
Session 2: Half			3.73	.013	H1	1400 (746)	H2	1291 (710)
Session * Type			9.07	<.001				
Session 1: Type			26.20	<.001				
Pairwise comparisons:	T Mm	vs C Mm		.007	Tone Mismatch			2019 (894)
		vs Match		<.001	Vowel Mismatch			1861 (835)
	V Mm	vs Match		<.001	Cons. Mismatch			1733 (823)
	C Mm	vs Match		<.001	Match			2316 (781)
Session 2: Type			37.96	<.001				
Pairwise comparisons:	T Mm	vs V Mm		<.001	Tone Mismatch			1444 (846)
		vs C Mm		<.001	Vowel Mismatch			1132 (750)
		vs Match		.001	Cons. Mismatch			1140 (676)
	V Mm	vs Match		<.001	Match			1666 (744)
	C Mm	vs Match		<.001				

Electrophysiology

Results for gRMS peak latency

Two gRMS peaks emerged in the data for both participant groups: ~180 ms (visual N1 latency) and ~370 ms (N400/LAN latency) after the picture onset. In the group average, the first peak was delayed for the TL1 group compared to the NTL1 group, cf. Figure 4. A 2-tailed independent samples t-test showed significant differences in average gRMS peak latency between the learner groups, $t(40) = -2.22$, $p = .032$. The TL1 group's first gRMS peak (M = 185 ms, SD = 12) was significantly delayed compared to the NTL1s' (M = 176 ms, SD = 13). Accordingly, the time windows for the first gRMS peak were 185-255 ms for the TL1 group and 176-246 ms for the NTL1 group (i.e., peak +70 ms). For the second gRMS peak (M_{TL1} = 373 ms, SD = 49, M_{NTL1} = 379 ms, SD = 44), there were no significant differences in timing. Hence, the time window submitted for the second gRMS peak was defined as 370-570 ms (i.e., ~peak +200 ms).

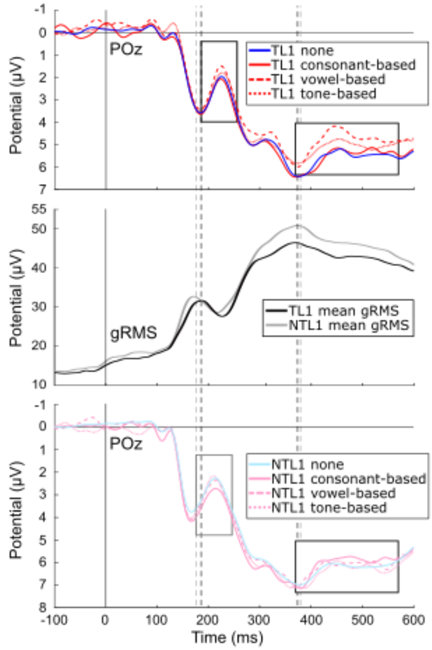


Figure 4. Middle: Latency of gRMS peaks for both learner groups, both sessions and all conditions averaged. Top: ERPs averaged over both sessions separately for all four experimental conditions in the tonal learner group at posterior electrode POz. Bottom: ERPs averaged over both sessions separately for all four experimental conditions in the non-tonal learner group at POz. Vertical dashed lines show the group's average gRMS peaks calculated from the subjects' individual average peak in each condition. Black = TL1, grey = NTL1.

Results for the ERP data

At N1 latency, the permutation analysis did not find any electrode clusters with different amplitudes for any mismatch condition or interaction. At N400/LAN latency, the permutation analysis produced a negative, centroposterior cluster (CP3, CP1, CP2, CP4, P1, P2), p = .005, indicative of an N400 effect, as well two positive, anterior clusters: Fpz, F7, F5, F3, F4, p = .021, and F1, Fz, FC1, p = .034 (cf. Figure 5 B) for the difference between consonant mismatch and matched pictures. The permutation test also identified a difference between vowel mismatch and matched pictures which manifested as a negative cluster at centroposterior electrodes (C6, CP3, CP4, CP6, TP8, P7, P5, P3, P1, Pz, P2, P4, P8, P07, P08, Oz), p = .001, again indicating an N400 response, as well as three positive clusters at anterior electrodes: AF4, F1, Fz, F2, FC3, FC1, FC2, C1, p = .002, F3, F4, Fpz, AF7, AF8, F5, p = .004, and Fp1, Fp2, F6, p = .035 (cf. Figure 5 B). Finally, in the comparison of tone mismatched and matched picture conditions, the permutation analysis did not find any differences for both groups collectively. The interaction with Learner Group for tone mismatch compared to matched pictures produced an N400-related centroposterior cluster (P1, Pz, P2), p = .043, (cf. Figure 5 B) where amplitudes were significantly more negative in the TL1 group than the NTL1 group. There were no interactions with Learner Group in the other conditions. Interactions with the temporal factors Session and Half or combinations thereof did not produce any significant clusters for either of the mismatch types.

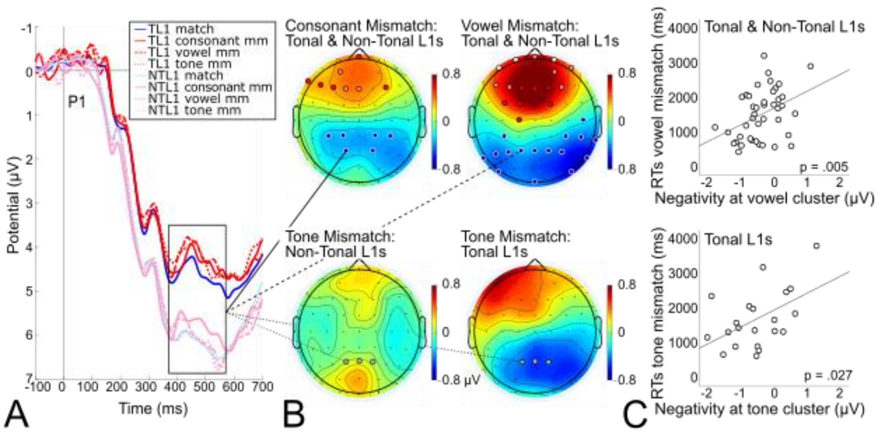


Figure 5. A. ERPs for the two participant groups and all conditions at posterior electrode P1. B. Subtraction topographies for the different mismatch-match conditions. Significant negative clusters indicated with blue dots, positive clusters range from dark red (most significant) to white (least significant), and interaction clusters with grey dots. C. Scatterplots for the correlation of negativity at N400 clusters and the response times for vowel mismatch in both groups, as well as negativity at N400 clusters and RTs for tone mismatch in the TL1 group.

Correlation behavioural and neurophysiological results

When testing for correlations between response times and ERP effects for error processing (i.e., N400), we found no significant correlations in the consonant condition. In the vowel condition, there was a significant positive correlation between vowel N400 and response times to vowel mismatches, $r(40) = .370$, p = .016: The larger the N400 effect for vowel, the faster the response times to vowel mismatches (cf. Figure 5 C). For the tone N400 in the TL1 group, we similarly observed a positive correlation between tone N400 and response times to tone mismatches, $r(19) = .452$, p = .040. (For the non-tonal learners, in comparison, neural responses to tone mismatches and response times were not correlated: $r(19) = .033$, p = .889).

Discussion

The present study investigated the processing of lexicosemantic and morphosyntactic errors in initial second language acquisition and the role of transfer in this process. In a word-picture association paradigm, learners acquired novel words with grammatical tone and, as a control, grammatical vowel change. All participants were familiar with systematic, grammatically distinctive vowel alterations from their L1. With respect to grammatical tone, however, the learners differed, such that they either were (TL1s) or were not (NTL1s) familiar with grammatically meaningful tone from their L1. This created the possibility of differential L1-L2 phonological transfer in the novel words. We occasionally presented pictures that were mismatched with the different phonological features of the preceding auditory words in the word-picture association paradigm. By doing so, we introduced lexicosemantic and morphosyntactic errors during the first four hours of acquisition. We expected to find neurophysiological responses related to lexicosemantic (N400) or morphosyntactic (E/LAN, P600) processing, respectively, for all phonological mismatch conditions in the TL1 learners. We anticipated grammatical ERP components in this learner group even though grammatical processing usually emerges relatively late in L2 learners. Acquisition conditions in the present study were optimised by allowing for complete L1-L2 transfer and through a strong focus on grammar. Furthermore, as mentioned in the introduction, we knew from previous studies that learners can deduce the grammatical content of the pictures in the word-picture association paradigm. We therefore assumed that grammatical ERP components, most likely a P600 and possibly a LAN or ELAN, would emerge for grammatically mismatched pictures in the tonal learners. For the NTL1 group, on the other hand, we expected reduced or missing ERP effects and possibly poorer behavioural performance for errors which were based on mismatches across the tone dimension but similar processing to the TL1 group for the other mismatch conditions.

Differential processing demands in beginner learners: Posterior N1

Somewhat unexpectedly, an early gRMS peak pointed at an important role of the visual perception-related posterior N1. The N1 (or N170), with a posterior distribution for visual input, is an exogenous response which is sensitive to task demands (Callaway & Halliday, 1982; Nash & Williams, 1982). Like other ERP components, the N1 is usually analysed with respect to amplitude differences between conditions or groups. The posterior N1 in the present study, however, showed timing rather than amplitude differences: The TL1 group's N1 was delayed in relation to the NTL1 group's N1. Such differences in N1 timing, although rare, have been attested in the literature and are generally associated with differentially effortful processing. As an example from the auditory domain, for instance, word recognition tasks in children with specific language impairment have elicited a delayed N1 compared to the N1 response from a matched control group of normally developing children (Malins, Desroches, Robertson, Newman, Archibald & Joanisse, 2013). With visual input, similar results have been reported when stimuli were processed under differential task demands. Children diagnosed with ADHD, for example, had a delayed N1 response to light flashes during an active detection task compared to a passive viewing (Callaway & Halliday, 1982). Similarly, healthy adults have shown a delayed N1 when tasked with stimulus identification rather than simple stimulus detection for moving visual stimuli (Fort, Besle, Giard & Pernier, 2005). In line with those studies, we believe that the N1 latency in the present study is related to differential cognitive demands for picture perception and identification in the two learner groups and that this is based on top-down influences. From the word-picture association, the tonal learners rapidly internalise the words' three phonological features (tone, vowel, and consonant) and their respective meaning. The activation of these features during word processing likely pre-activates the associated pictorial properties (number, gender, and profession) in the TL1 learners prior to the emergence of the picture. Thus, when the picture referent appears in the input, the TL1 learners are faced with the difficult task of pre-consciously checking and identifying the image against the three pre-activated categories. The NTL1 learners, on the other hand, have presumably only internalised two phonological categories, i.e., vowel and consonant, and their corresponding meaning, which influence their pre-conscious perception and the picture identification task. We believe that the automatic access to a greater number of pictorial properties, through pre-activation, increased automatic picture identification difficulty in the non-tonal learners which manifested in a delayed posterior N1. Interestingly, there were no differences in either N1 latency or amplitude between mismatching or matching pictures in either participant group, suggesting that validity does not affect processing at this latency. Importantly, the N1 results show that the absence or presence of positive L1-L2 transfer in auditory words has a strong influence on the pre-attentive processing of upcoming pictorial referents. Hence, transfer not only affects the processing of the transfer-facilitated words themselves but also allows for pre-activation of closely related items in the immediate context via top-down processes which influences perception. This finding calls for more detailed investigations into the generalisability and scope of transfer-facilitated processing effects. Can a comparable N1 latency effect also be found for written word form identification and possibly even in the auditory modality? If so, how close to each other do the associated elements need to be? Can positive transfer effects on the first element in a non-adjacent dependency, for instance, influence the pre-attentive perception and identification of the second element further down the language stream? Addressing these and related questions could provide important insights for a more comprehensive understanding of the N1 in the perceptive processing of pre-activated elements in language contexts.

Mismatch responses in beginner learners: N400 as a response to all error types

For mismatch processing, we anticipated differential results depending on mismatch type. For semantic errors, which were contained in mismatches between the words' consonants and the profession depicted in the associated picture, we expected an N400 response. For morphosyntactic errors, which were elicited by mismatches of picture properties related to the words' vowels or tones, on the other hand, we expected to see morphosyntactic ERP components. The vowels and tones had clear and easily abstractable morphosyntactic content and L1-L2 transfer was possible. For comparable circumstances, previous studies on artificial acquisition rapidly elicited at the very least a P600 (e.g., Batterink & Neville, 2013; Havas et al., 2020) and sometimes even additional early frontal negativities (de Diego-Balaguer et al., 2007). However, we find no evidence of a P600-related late positivity in response to the morphosyntactic mismatches in the present study. Although surprising, the lack of a P600 might be explained by the fact that mismatches were between novel spoken words and pictorial referents. While N400 and LAN effects have previously been observed for pictures in linguistic contexts (e.g., Federmeier & Kutas, 2001; Wicha et al., 2003), a P600 was never reported in those studies. It is thus possible that the absence of the P600 in the present study is simply due to the mismatch format. Instead of a P600, we found an earlier negativity, as indicated by a prominent gRMS peak at ~370 ms. This negativity was related to an N400 response and was found consistently for all mismatch types, but only when L1-L2 transfer was possible. The emergence of an N400 rather than the expected LAN may be an unintended result of the use of visual referents. The pictorial form of the mismatches may automatically have induced semantic processing. We assume that learners are aware of 'gender' and 'number' categories in a word-picture association paradigm, as the related visual cues in the pictures can install grammatical processing in the words (Gosselke Berthelsen et al., 2020). However, this does not necessarily entail that the same pictorial features could not in theory be processed semantically when they are presented as part of the pictures, especially in L2 contexts. That is to say, the abstracted, morphosyntactic label 'feminine' for the words could be processed concretely as the semantic feature 'female' during picture processing using visual features. For number this would correspond to a grammatical vs numeric number distinction between words and pictures such as 'plural' vs 'three'. The pictorial form of the mismatches is thus an entirely possible explanation of the absence of P600 and LAN effects.

An alternative explanation for the lack of morphosyntactic responses in the present data might be categorical differences between canonical processing and error processing in L2 learners. In a direct comparison of results elicited with the present word-picture association learning paradigm, we find an ELAN, LAN and P600 for canonical words but a N400 for mismatched pictures. This shows that learners can use the present paradigm to produce a full range of grammatical components for newly acquired words. These are elicited after just 20 minutes of acquisition in the comparison of learned words and meaningless pseudowords and thus for canonical processing (Gosselke Berthelsen et al., 2020). For errors in the word-picture pairs, in contrast, we do not observe morphosyntactic processing although this should in theory have been possible. Previous literature does find quickly emerging morphosyntactic components even in error contexts (Batterink & Neville, 2013; de Diego-Balaguer et al., 2007; Havas et al., 2020). Those studies, however, either had a much stronger, solitary focus on syntax (Batterink & Neville, 2013; de Diego-Balaguer et al., 2007), or a considerably less complex morphosyntactic system (Havas et al., 2017). The learners in the present study, instead, produced an initial L2 default response to all types of L2 errors: an N400 (Osterhout, Poliakov, Inoue, McLaughlin, Valentine, Pitkanen, Frenck-Mestre & Hirschensohn, 2008). This default response suggets that differentiations of mismatches as lexicosemantic or morphosyntactic do not take place. This differentiation might generally be of little importance to beginner learners. Error detection arguably takes precedence over error classification. In fact, this might be an important factor to keep in mind for our interpretation of grammar processing in second language learners. Traditionally, grammar processing is measured with the help of violation paradigms. The present results hint that this very tradition might potentially skew the picture. Learners might not in fact be incapable of grammar processing per se but instead simply be indifferent to different error types. They might then respond only to the unexpectedness of an erroneous item and the effect that this has on utterance-level meaning, something which would elicit an N400. Therefore, the late emergence of full-fledged grammatical processing of errors in both natural (e.g., Gillon Dowens et al., 2010; Tanner et al., 2012) and artificial second language acquisition (cf. e.g., Tanner et al., 2012; Morgan-Short et al., 2012) might be related to the fact that the learners' L2 processing resources are initially exhausted by canonical processing and the detection but not classification of errors. In summary, the present findings suggest a need for carefully evaluating the use of violation paradigms as measurements of L2 proficiency in beginner learners and, more generally, grammar processing in second language contexts. Future studies should more closely investigate potential categorical differences between canonical processing and error processing in L2 learners. Also, to untangle the contribution of error processing and picture processing from the current findings, it is essential that future studies test the processing of grammar errors contained in auditory words rather than pictures. This could even be achieved using the same word-picture association paradigm as in the present study but including intermitted reverse test pairs (i.e., picture-word) after every learning cycle where matches and mismatches are then contained in the auditory words.

The second main finding in the mismatch processing responses is the different processing of tone-mismatched picture referents between learner groups. This emphasises once more the importance of L1-L2 similarity in initial L2 acquisition: Only mismatches based on familiar phonological features elicited a mismatch response in the learners. This is presumably related to the inability to pre-consciously process unfamiliar phonological elements; a realisation that is well-attested in the rapid learning literature (e.g., Gosselke et al., 2020; Kimppa et al., 2015; Silva et al., 2019). The non-tonal learners cannot pre-attentively access the word's tone or its content, which in turn likely inhibits automatic pre-activation of the tone-related picture features. As there are no pre-activated features for the tone condition for the NTL1s, the appearance of a mismatch in the pictures will not lead to an expectancy-based mismatch, hence preventing the emergence of an N400. Notice, however, that the lack of an expectancy-related mismatch response in the NTL1 group does not entail that mismatches were not detected. Behavioural results show no significant differences between the tonal and non-tonal learners in tone mismatch detection accuracy (MTL1 = 68.9%; MNTL1 = 62.6%) or response times (MTL1 = 1724 ms; MNTL1 = 1853 ms), replicating findings which show that transfer affects online but not offline processing in L2 learners (e.g., Andersson et al., 2019). For features where transfer is not possible, learners presumably rely on later, more consciously evaluative processing to detect mismatches. This contrasts with pre-activation-induced mismatch detection for transferable features, which is likely initiated during pre-conscious processing of the feature and its associated meaning. Interestingly, the additive nature of the grammatical tone entailed that the non-tonal learners could process the remainder of the word, i.e., the consonant and vowel content, independently of the tone. We could, therefore, observe an N400 for errors based on mismatches across the vowel and consonant dimension even in the NTL1 group. The transfer-related effect of the words' phonological features on the processing of mismatches in pictures again shows that transfer affects not only the word itself but also the processing of tightly associated referents, demonstrating the overarching influence of L1-L2 similarity in SLA. As argued for the N1, above, further studies are needed to determine the potential reach of such transfer-based top-down effects on the processing of upcoming words. It would be useful to determine, for instance, whether only the immediately following element can be top-down mediated or whether this effect extends to any closely associated element, possibly with a number of intervening constituents.

Finally, in accordance with previous findings on the rapidness of initial L2 acquisition (e.g., Cunillera et al., 2009; de Diego-Balaguer et al., 2007; Dittinger et al., 2019; Gullberg et al., 2010; Gosselke et al., 2020; Kimppa et al., 2015), the N400 for mismatched pictures manifested already within 1 hour, i.e., within 15 learning cycles and 15 mismatch trials per condition. At this point, mismatch processing had manifested to such a degree that further changes over time were undetectable. Indeed, the effect was likely present already after 30 minutes. We see great improvements in the behavioural responses within the first session and know from previous literature that neurophysiological changes tend to precede behavioural ones (e.g., McLaughlin et al., 2010). A targeted investigation into the onset of error processing in L2 acquisition would require a reliable way of studying ERP responses at smaller time intervals, for instance by means of testing a very large number or participants. Alternatively, future studies could use an acquisition paradigm with a considerably larger number of target words so that the number of error trials increases as a function of the total number of trials.

Behavioural responses to errors in initial learners

The two learner groups in the present study were virtually indistinguishable in their behavioural performance on mismatch detection. Although mean accuracy for tone mismatches was slightly lower in the non-tonal learner group and mean response times longer, between-group differences were not significant. The observed findings tie in nicely with previous literature showing that learners' offline responses are often unaffected by transfer (e.g., Andersson et al., 2020). Instead, facilitation effects based on L1-L2 similarity tend to manifest in online measures. Learners who cannot profit from positive L1-L2 transfer have to resort to different types of processing, for instance conscious rather than pre-conscious (cf. Gosselke Berthelsen et al., 2020), but reach the same learning outcome.

A secondary finding in the behavioural data in the present study was an overall poor acquisition outcome for tone. Of all mismatch conditions, tone had the lowest accuracy and slowest response times. The inferior learning performance is likely based on two factors: a slight functional dissimilarity in the L1 and L2 tone in this study and a general difficulty in the acquisition of L2 tone. The tonal participants' L1 tone is associated with grammar but not itself grammatically meaningful, unlike the L2 tone in the present study. This might have negatively impacted on the TL1s' tone acquisition. Besides, L2 tone is generally difficult to acquire and even distinguish. To this effect, Yang and Chan (2010) found that even highly advanced learners of Chinese performed poorly in the perceptual discrimination of tone. For the least accurate tone contrast in their study, discrimination accuracy was below 30% in the advanced learners, compared to nearly 100% in the native speaker control group. This considerable discrimination difference is a clear indicator of the general perceptual and in turn acquisitional challenges associated with L2 tone. Coupled with the slight L1-L2 tone dissimilarity, this readily explains the lower acquisition outcome for L2 tone even in the tonal learners.

Interestingly, however, while there were no between-group differences for mismatch detection in the behavioural variables, we found a correlation between behavioural factors and the N400 such that response times were faster the larger the N400. This correlation centrally emerged both for vowel mismatch and in the tonal learners also for tone mismatch. The emergence of the correlation in the tonal learners (and a clear lack of correlation in the NTL1s) strengthens the suggested impact of the words' phonological properties, via pre-activation of linguistic content, in the processing and detection of errors in picture referents. Those learners, who react most strongly to picture mismatches neurophysiologically (visible in larger N400 effects), are also fastest as identifying a corresponding error behaviourally. Both outcomes are likely mediated through expectation-related processes.

Conclusions

In the current study, we investigated how behavioural error detection and neurophysiological error processing proceeded in novice learners. We found that mismatches elicited by pictures in word-picture pairs uniformly elicited an N400 response, regardless of whether the mismatch was lexicosemantic or morphosyntactic in nature. We believe this to be indicative of the increased difficulty of error processing in second language learners, making the study of grammatical ERP responses in violation paradigms a questionable measure of L2 proficiency and processing. The N400 emerged well within an hour of acquisition and was correlated with behavioural responses. Importantly, we only observed an N400 when the error was based on a familiar phonological property in the preceding word, highlighting the importance of cross-linguistic transfer. Besides the N400, we found a visual perception-related posterior N1 for the pictures which was delayed for learners who could, via phonological transfer, internalise all three sound-meaning relationships and thus automatically pre-activate three rather than only two of the picture's associated visual properties.

References

Andersson, A., Sayehli, S., & Gullberg, M. (2019). Language background affects online word order processing in a second language but not offline. Bilingualism: Language and Cognition, 22(4), 802-825. doi:10.1017/S1366728918000573

Barber, H., & Carreiras, M. (2005). Grammatical gender and number agreement in Spanish: An ERP comparison. Journal of Cognitive Neuroscience, 17(1), 137-153. doi:10.1162/0898929052880101

Batterink, L., & Neville, H. (2013). Implicit and explicit second language training recruit common neural mechanisms for syntactic processing. Journal of Cognitive Neuroscience, 25(6), 936-951. doi:10.1162/jocn_a_00354.

Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. Journal of the Royal Statistical Society. Series B (Methodological), 57(1), 289-300. doi:10.2307/2346101

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

Bowden, H. W., Steinhauer, K., Sanz, C., & Ullman, M. T. (2013). Native-like brain processing of syntax can be attained by university foreign language learners. Neuropsychologia, 2492-2511. doi:10.1016/j.neuropsychologia.2013.09.004

Braun, B., & Johnson, E. K. (2011). Question or tone 2? How language experience and linguistic function guide pitch processing. Journal of Phonetics, 39, 585-594. doi:10.1016/j.wocn.2011.06.002

Callaway, E., & Halliday, R. (1982). The effect of attentional effort on visual evoked potential N1 latency. Psychiatry Research, 7(3), 299-308. doi:10.1016/0165-1781(82)90066-X

Cunillera, T., Càmara, E., Toro, J. M., Marco-Pallares, J., Sebasián-Galles, N., Ortiz, H., . . . Rodríguez-Fornells, A. (2009). Time course and functional neuroanatomy of speech segmentation in adults. 48(3), 541-553. doi:/10.1016/j.neuroimage.2009.06.069

Davidson, D. J., & Indefrey, P. (2009). Plasticity of grammatical recursion in German learners of Dutch. Language and Cognitive Processes, 24(9), 1335-1369. doi:10.1080/01690960902981883

de Diego-Balaguer, R., Rodríguez-Fornells, A., & Bacoud-Lévi, A.-C. (2015). Prosodic cues enhance rule learning by changing speech segmentation mechanisms. Frontiers in Psychology, 6(1478). doi:10.3389/fpsyg.2015.01478

de Diego-Balaguer, R., Toro, J. M., Rodríguez-Fornells, A., & Bachoud-Lévi, A.-C. (2007). Different neurophysiological mechanisms underlying word and rule extraction from speech. PLos ONE 2, 11. e1175. doi:10.1371/journal.pone.0001175

Dittinger, E., Scherer, J., Jäncke, L., Besson, M., & Elmer, S. (2019). Testing the influence of musical expertise on novel word learning across the lifespan using a cross-sectional approach in children, younger adults and older adults. Brain and Language, 198(104678). doi:10.1016/j.bandl.2019.104678

Federmeier, K. D., & Kutas, M. (2001). Meaning and modality: Influences of context, semantic memory organization, and perceptual predictability on picture processing. Journal of Experimental Psychology: Learning, Memory, and Cognition, 27(1), 202-224. doi:10.103I77////0022778-7393.27.1.202

Fort, A., Besle, J., Giard, M.-H., & Pernier, J. (2005). Task-dependent activation latency in human visual extrastriate cortex. Neuroscience Letters, 379, 144-148. doi:10.1016/j.neulet.2004.12.076

Foucart, A., & Frenck-Mestre, C. (2011). Grammatical gender processing in L2: Electrophysiological evidence of the effect of L1-L2 syntactic similarity. Bilingualism: Language and Cognition, 14(3), 379-399. doi:10.1017/S136672891000012X

Foucart, A., & Frenck-Mestre, C. (2012). Can late L2 learners acquire new grammatical features? Evidence from ERPs and eye-tracking. Journal of Memory and Language, 66, 226-248. doi:10.1016/j.jml.2011.07.007

François, C., Cunillera, T., Garcia, E., Laine, M., & Rodriguez-Fornells, A. (2017). Neurophysiological evidence for the interplay of speech segmentation and word-referent mapping during novel word learning. Neuropsychologia, 56-67. doi:10.1016/j.neuropsychologia.2016.10.006

Friederici, A. D. (2002). Towards a neural basis of auditory sentence processing. TRENDS in Cognitive Science, 6(2), 78-84. doi:10.1016/S1364-6613(00)01839-8

Friederici, A. D., & Weissenborn, J. (2007). Mapping sentence form onto meaning: The syntax-semantic interface. Brain Research, 1146, 50-58. doi:10.1016/j.brainres.2006.08.038

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

Ganis, G., Kutas, M., & Sereno, M. I. (1996). The search for "common sense": An electrophysiological study of the comprehension of words and pictures in reading. Journal of Cognitive Neuroscience, 8(2), 89-106. doi: 10.1162/jocn.1996.8.2.89

Gillon Dowens, M., Guo, T., Guo, J., Barber, H., & Carreiras, M. (2011). Gender and number processing in Chinese learners of Spanish - Evidence from Event Related Potentials. Neuropsychologia, 49, 1651-1659. doi:10.1016/j.neuropsychologia.2011.02.034

Gillon Dowens, M., Vergara, M., Barber, H. A., & Carreiras, M. (2010). Morphosyntactic processing in late second-language learners. Journal of Cognitive Neuroscience, 22(8), 1870-1887. doi:10.1162/jocn.2009.21304

Gosselke Berthelsen, S., Horne, M., Brännström, K. J., Shtyrov, Y., & Roll, M. (2018). Neural processing of morphosyntactic tonal cues in second-language learners. Journal of Neurolinguistics, 45, 60-78. doi:10.1016/j.jneuroling.2017.09.001

Gosselke Berthelsen, S., Horne, M., Shtyrov, Y., & Roll, M. (2020). Different neural mechanisms for rapid acquisition of words with grammatical tone in learners from tonal and non-tonal backgrounds: ERP evidence. Brain Research, 1729(146614). doi:10.1016/j.brainres.2019.146614

Grey, S., Sanz, C., Morgan-Short, K., & Ullman, M. T. (2018). Bilingual and monolingual adults learning an additional language: ERPs reveal differences in syntactic processing. Bilingualism: Language and Cognition, 21(5), 970-994. doi:10.1017/S1366728917000426

Guarjardo, L. F., & Wicha, N. Y. (2014). Morphosyntax can modulate the N400 component: Event related poentials to gender-marked post-nominal adjectives. NeuroImage, 91, 262-272. doi:10.1016/j.neuroimage.2013.09.077

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

Hallé, P. A., Chang, Y.-C., & Best, C. T. (2004). Identification and discrimination of Mandarin Chinese tones by Mandarin Chinese vs. French listeners. Journal of Phonetics, 32, 395-42. doi: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. doi:10.1016/j.neuropsychologia.2016.01.001

Hautus, M. J. (1995). Corrections for extreme proportions and their biasing effects on estimated values of d'. Behavior Research Methods, Instruments, & Computers, 27, 46-51. doi:10.3758/BF03203619

Havas, V., Laine, M., & Rodriguez-Fornells, A. (2017). Brain signatures of early lexical and morphological learning of a new language. Neuropsychologia, 101, 47-56. doi:10.1016/j.neuropsychologia.2017.04.005

Havas, V., Taylor, J., Vaquero, L., de Diego-Balaguer, R., Rodríguez-Fornells, A., & Davis, M. H. (2018). Semantic and phonological schema influence spoken word learning and overnight consolidation. Quarterly Journal of Experimental Psychology, 71(6), 1469-1481. doi:10.1080/17470218.2017.1329325

Hed, A., Schremm, A., Horne, M., & Roll, M. (2019). Neural correlates of second language acquisition of tone-grammar associations. The Mental Lexicon, 14(1), 98-123. doi:10.1075/ml.17018.hed

Hollingshead, A. B. (1975). Four factor index of social status. Unpublished manuscript, Yale University, New Haven, CT.

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.

Jung, T.-P., Makeig, S., Humphries, C., Lee, T.-W., McKeown, M. J., Iragui, V., & Sejnowski, T. J. (2000). Removing electroencephalographic artifacts by blind source separation. Psychophysiology, 37, 136-178. doi:10.1111/1469-8986.3720163

Kim, A., & Osterhout, L. (2005). The independence of combinatory semantic processing: Evidence from event-related potentials. Journal of Memory and Language, 52(2), 205-225. doi:10.1016/j.jml.2004.10.002

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. doi:10.1016/j.neuroimage.2015.05.098

Krott, A., & Lebib, R. (2013). Electrophysiological evidence for a neural substrate of morphological rule application in correct wordforms. Brain Research, 1496, 70–83. doi:10.1016/j.brainres.2012.12.012

Kuperberg, G. R., Holcomb, P. J., Stinikova, T., Greve, D., Dale, A. M., & Caplan, D. (2003). Distinct patterns of neural modulation during the processing of conceptual and syntactic anomalies. Journal of Cognitive Neuroscience, 15(2), 272-293. doi:10.1162/089892903321208204

Kutas, M., & Federmeier, K. D. (2014). Thirty years and counting: Finding meaning in the N400 component of the event related brain potential (ERP). Annual Review of Psychology, 62, 621-647. doi: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. doi:10.1126/science.7350657

Kutas, M., & van Petten, C. (1990). Electrophysiological perspectives on comprehending written language. New Trends and Advanced Techniques in Clinical Neurophysiology, 41, 155-167. doi:10.1016/b978-0-444-81352-7.50020-0

Lehmann, D., & Skrandies, W. (1980). Reference-free identification of components of checkerboard-evoked multichannel potential fields. Electroencephalography and Clinical Neurophysiology, 48(6), 609-621. doi:10.1016/0013-4694(80)90419-8

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. doi:10.1016/j.cortex.2016.06.020

Lozano-Argüelles, C., Sagarra, N., & Casillas, J. V. (2019). Slowly but surely: Interpreting facilitates L2 morphological anticipation based on suprasegmental and segmental information. Bilingualism: Language and Cognition, 1-11. doi:10.1017/S1366728919000634

Malins, J. G., Desroches, A. S., Robertson, E. K., Lynn, N. R., Archibald, L. M., & Joanisse, M. F. (2013). ERPs reveal the temporal dynamics of auditory word recognition in specific language impairment. Developmental Cognitive Neuroscience, 5, 134-148. doi:10.1016/j.dcn.2013.02.005

Maris, E., & Oostenveld, R. (2007). Nonparametric statistical testing of EEG- and MEG-data. Journal of Neuroscience Methods, 164, 177-190. doi:10.1016/j.jneumeth.2007.03.024

McKean, C., Letts, C., & Howard, D. (2013). Functional reorganization in the developing lexicon: separable and changing influences of lexical and phonological variables on children's fast-mapping. Journal of Child Language, 40, 307-335. doi:10.1017/S0305000911000444

McLaughlin, J., Osterhout, L., & Kim, A. (2004). Neural correlates of second-language word learning: minimal instruction produces rapid change. Nature Neuroscience, 7(7), 703-704. doi:10.1038/nn1264

McLaughlin, J., Tanner, D., Pitkänen, I., Frenck-Mestre, C., Inoue, K., Valentine, G., & Osterhout, L. (2010). Brain potentials reveal discrete stages of L2 grammatical learning. Language Learning, 60(Suppl. 2), 123-150. doi:10.1111/j.1467-9922.2010.00604.x

Molinaro, N., Barber, H. A., Caffarara, S., & Carreiras, M. (2014). On the left anterior negativity (LAN): The case of morphosyntactic agreement. Cortex, 1-4. doi:10.1016/j.cortex.2014.06.009

Morgan-Short, K., Steinhauer, K., Sanz, C., & Ullman, M. T. (2012). Explicit and implicit second language training differentially affect the achievement of native-like brain activation patterns. Journal of Cognitive Neuroscience, 24(4), 993-947. doi:10.1162/jocn_a_00119

Mueller, J. L., Hahne, A., Fujii, Y., & Friederici, A. D. (2005). Native and nonnative speakers' processing of a miniature version of Japanese as revealed by ERPs. Journal of Cognitive Neuroscience, 17(8), 1229-1244. doi:10.1162/0898929055002463

Mueller, J. L., Hirotani, M., & Friederici, A. D. (2007). ERP evidence for different strategies in the processing of case markers in native speakers and non-native learners. BMC Neuroscience, 8(18). doi:10.1162/0898929055002436

Nash, A. J., & Williams, C. S. (1982). Effects of preparatory set and task demands on auditory event-related potentials. Biological Psychology, 15, 15-31. doi:10.1016/0301-0511(82)90028-X

Neville, H., Nicol, J. L., Barss, A., Forster, K. I., & Garret, M. F. (1991). Syntactically based sentence processing classes: Evidence from event-related potentials. Journal of Cognitive Neuroscience, 3(2), 151-165. doi:10.1162/jocn.1991.3.2.151

Nigam, A., Hoffman, J. E., & Simons, R. F. (1992). N400 to semantically anomalous pictures and words. Journal of Cognitive Neuroscience, 4(1), 15-22. doi:10.1162/jocn.1992.4.1.15

Ojima, S., Nakata, H., & Kakigi, R. (2005). An ERP study of second language learning after childhood: Effects of proficiency. Journal of Cognitive Neuroscience, 17(8), 1212-1228. doi:10.1162/0898929055002436

Osterhout, L., & Holcomb, P. J. (1992). Event-related brain potentials elicited by syntactic anomaly. Journal of Memory and Language, 31, 785-806. doi:10.1016/0749-596X(92)90039-Z

Osterhout, L., Poliakov, A., Inoue, K., McLaughlin, J., Valentine, G., Pitkanen, I., . . . Hirschensohn, J. (2008). Second-language learning and changes in the brain. Journal of Neurolinguistics, 21(6), 509-521. doi:10.1016/j.jneuroling.2008.01.001

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 children as revealed by neuromagnetic brain response dynamics. NeuroImage, 155, 450-459. doi:10.1016/j.neuroimage.2017.03.066

Pelzl, E., Lau, E. F., Guo, T., & DeKeyser, R. (2019). Advanced second language learners' perception of lexical tone contrasts. Studies in Second Language Acquisition, 41, 59-86. doi:10.1017/S0272263117000444

Pratarelli, M. E. (1994). Semantic processing of pictures and spoken words: Evidence from event-related brain potentials. Brain and Cognition, 24, 137-157. doi:10.1006/brcg.1994.1008

Roll, M., Gosselke, S., Lindgren, M., & Horne, M. (2013). Time-driven effects on processing grammatical agreement. Frontiers in Psychology, 4(1004), 1-8. doi:10.3389/fpsyg.2013.01004

Roll, M., Horne, M., & Lindgren, M. (2007). Object shift and event-related brain potentials. Journal of Neurolinguistics, 20(6), 462-481. doi:10.1016/j.jneuroling.2007.04.001

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 & Language, 150, 14-21. doi:10.1016/j.bandl.2015.07.009

Rossi, S., Gugler, M. F., Friederici, A. D., & Hahne, A. (2006). The impact of proficiency on syntactic second-langauge processing of German and Italian: Evidence from event-related potentials. Journal of Cognitive Neuroscience, 18(12), 2030-2048. doi:10.1162/jocn.2006.18.12.2030

Sagarra, N., & Casillas, J. V. (2018). Suprasegmental information cues morphological anticipation during L1/L2 lexical access. Journal of Second Language Studies, 1(1), 31-59. doi:10.1075/jsls.17026.sag

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. doi:10.1016/j.jneuroling.2019.01.006

Shen, G., & Froud, K. (2019). Electrophysiological correlates of categorical perception of lexical tones by English learners of Mandarin Chinese: an ERP study. Bilingualism: Language and Cognition, 22(2), 253-265. doi:10.1017/S136672891800038X

Silva, S., Vigário, M., Leone Fernandez, B., Jerónimo, R., Alter, K., & Frota, S. (2019). The sense of sounds: Brain responses to phonotactic frequency, phonological grammar and lexical meaning. frontiers in Psychology, 10(681). doi:10.3389/fpsyg.2019.00681

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

Söderström, P., Horne, M., & Roll, M. (2016). Stem tones pre-activate suffixes in the brain. Journal of Psycholinguistic Research, 1-10. doi:10.1007/s10936-016-9434-2

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. doi:10.1093/applin/amu028

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

Tanner, D., McLaughlin, J., Herschensohn, J., & Osterhout, L. (2012). Individual differences reveal stages of L2 grammatical acquisition: ERP evidence. Bilingualism: Language and Cognition, 16, 367-382. doi:10.1017/S1366728912000302

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. doi:10.1017/S0272263105050102

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. doi:10.3758/BF03192720

Wayland, R. P., & Li, B. (2008). Effects of two training procedures in cross-language perception of tones. Journal of Phonetics, 36, 250-267. doi:10.1016/j.wocn.2007.06.004

Wicha, N. Y., Moreno, E. M., & Kutas, M. (2003). Expecting gender: An event related brain potential study on the role of grammatical gender in comprehending line drawings within a written sentence in Spanish. Cortex, 39, 483-508. doi:10.1016/S0010-9452(08)70260-0

Williams, S. M. (2013). A major revision of the Edinburgh Handedness Inventory. doi: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.

Yang, J., & Li, P. (2019). Mechanisms for auditory perception: A neurocognitive study of second language learning of Mandarin Chinese. Brain Sciences, 9(139). doi:10.3390/brainsci9060139

Yu, K., Li, L., Chen, Y., Zhou, Y., Wang, R., Zhang, Y., & Li, P. (2019). Effects of native language experience on Mandarin lexical tone processing in proficient second language learners. Psychophysiology, 56:e13448. doi:10.1111/psyp.13448

Appendix A: Overview of picture stimuli

Table A. Examples of control stimuli used in combination with control words

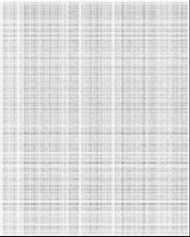
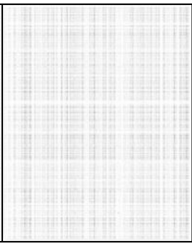
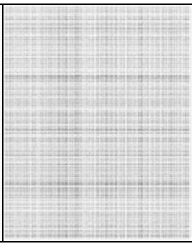
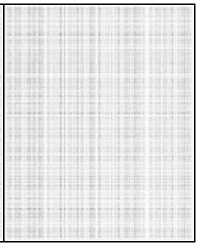
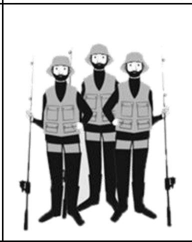
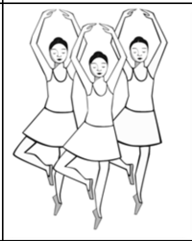
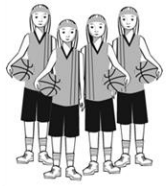
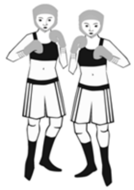
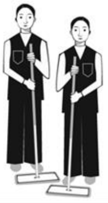
Control stimuli	no meaning			
no meaning				

Table B. Full list of target stimuli used to assign meaning to target words

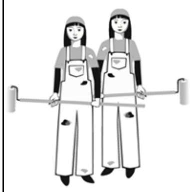
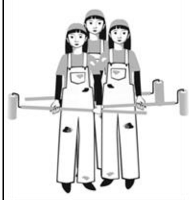
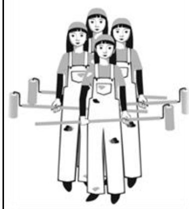
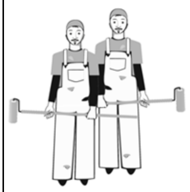
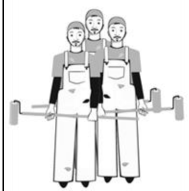
Profession		Singular	Plural		
Angler	Female (feminine)				
	Male (masculine)				
Ballet dancer	Female (feminine)				
	Male (masculine)				

Profession		Singular	Plural		
Basket-baller	Female (feminine)				
	Male (masculine)				
Boxer	Female (feminine)				
	Male (masculine)				
Chemist	Female (feminine)				
	Male (masculine)				

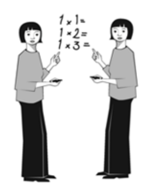
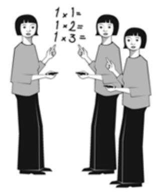
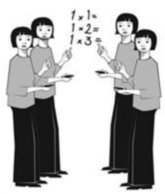
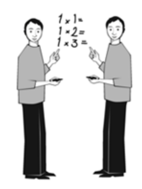
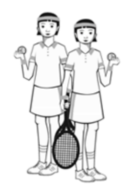
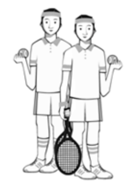
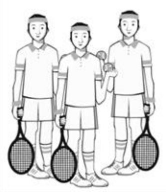
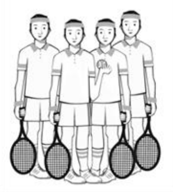
Profession		Singular	Plural		
Cleaner	Female (feminine)				
	Male (masculine)				
Cook	Female (feminine)				
	Male (masculine)				
Doctor	Female (feminine)				
	Male (masculine)				

Profession		Singular	Plural		
Fire fighter	Female (feminine)				
	Male (masculine)				
Flautist	Female (feminine)				
	Male (masculine)				
Flight attendant	Female (feminine)				
	Male (masculine)				

Profession		Singular	Plural		
Gardener	Female (feminine)				
	Male (masculine)				
Hair dresser	Female (feminine)				
	Male (masculine)				
Journalist	Female (feminine)				
	Male (masculine)				

Profession		Singular	Plural		
Painter	Female (feminine)				
	Male (masculine)				
Photo-grapher	Female (feminine)				
	Male (masculine)				
Police officer	Female (feminine)				
	Male (masculine)				

Profession		Singular	Plural		
Race driver	Female (feminine)				
	Male (masculine)				
Singer	Female (feminine)				
	Male (masculine)				
Tailor	Female (feminine)				
	Male (masculine)				

Profession		Singular	Plural		
Teacher	Female (feminine)				
	Male (masculine)				
Tennis player	Female (feminine)				
	Male (masculine)				
Violinist	Female (feminine)				
	Male (masculine)				

Profession		Singular	Plural		
Waiter	Female (feminine)				
	Male (masculine)				

Paper III

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 ($hatt_1en$, hat-the) rather than a plural suffix (*$hatt_1ar$, *hats) or a productive compound (*$hatt_1hylla$, *hat rack). Plural *-ar* and compounds instead have to be preceded by a high tone (i.e. accent 2 → $hatt_2ar$, hats; $hatt_2hylla$, 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

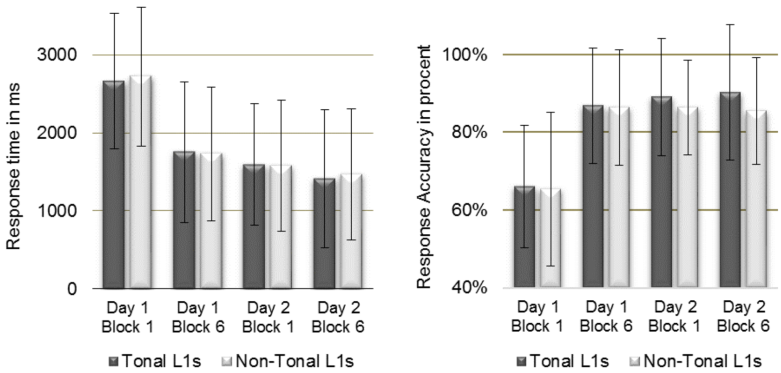


Fig. 1. Mean response time (left) and response accuracy (right) at the beginning and end of both days by group (bars represent standard deviations).

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.

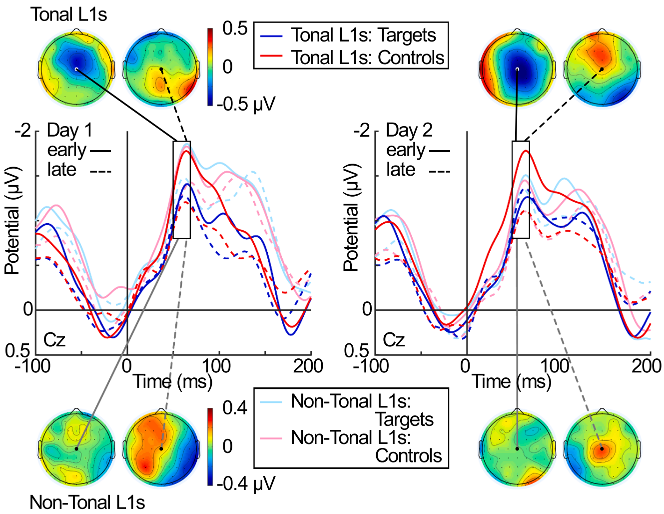


Fig. 2. Middle: ERPs from both groups at central electrode Cz for early and late trials on Day 1 (left) and Day 2 (right). Tonal L1s in dark colours, Non-Tonal L1s in pale colours. Early trials in continuous lines, late trials in dotted lines. Above and below: Topographical distributions for the effects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

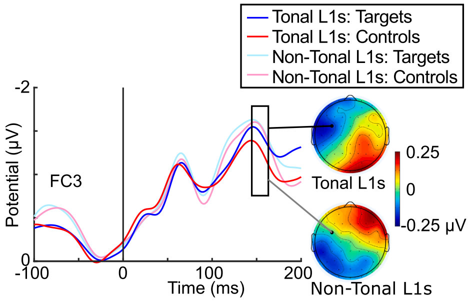


Fig. 3. Left: ERPs at left-frontal electrode FC3 for the two groups for all trials: Tonal L1s in dark colours, Non-Tonal L1s in pale colours. Right: Topographical distributions of target-control effects at ELAN latency (145–165 ms) for all trials in the Tonal L1 (above) and Non-Tonal L1 (below) group.

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

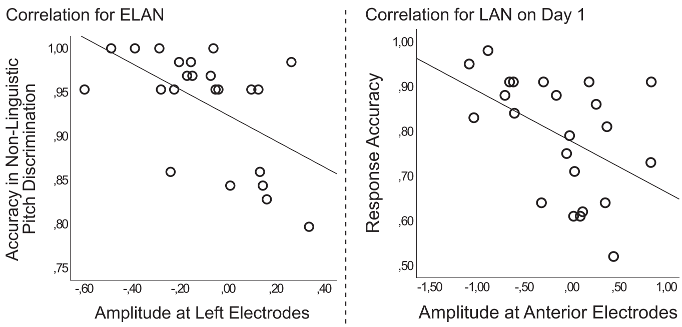


Fig. 4. Left: Scatterplot for the correlation of negativity at left lateral electrodes for targets-controls and Accuracy in Non-Linguistic Pitch Distinction in the Non-Tonal L1 group. Right: Scatterplot for the correlation of negativity at anterior electrodes for targets-controls on day 1 and Response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

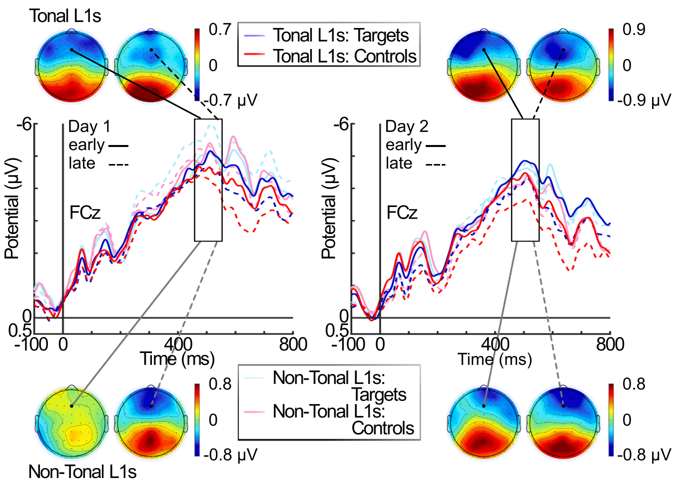


Fig. 5. Middle: ERPs from both groups at mid-frontal electrode FCz for early and late trials on Day 1 (left) and Day 2 (right). Tonal L1s in dark colours, Non-Tonal L1s in pale colours. Early trials in continuous lines, late trials in dotted lines. Above and below: Topographical distributions for the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

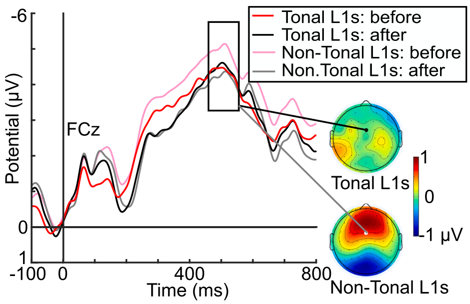


Fig. 6. Left: ERPs at anterior electrode FCz for the two groups for all trials: Tonal L1s in dark colours, Non-Tonal L1s in pale colours. Right: Topographical distributions of consolidation effects (after-before consolidation) at 460–560 ms for all trials in the Tonal L1 (above) and Non-Tonal L1 (below)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

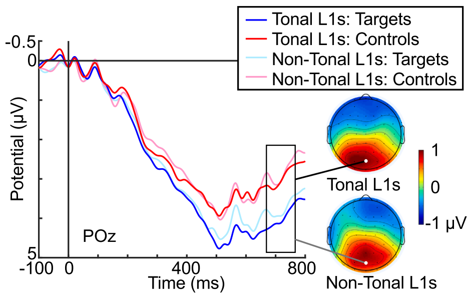


Fig. 7. Left: ERPs at posterior electrode POz for the two groups for all trials: Tonal L1s in dark colours, Non-Tonal L1s in pale colours. Right: Topographical distributions of target-control effects at P600 latency (670–770 ms) for all trials in the Tonal L1 (above) and Non-Tonal L1 (below) group.

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

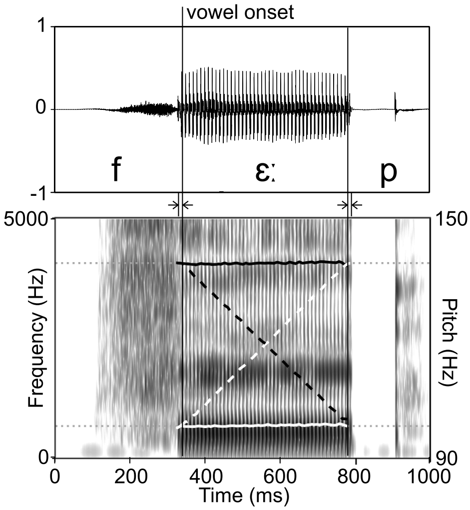


Fig. 8. Example of a stimulus word with acoustic waveform, spectrogram, and fundamental frequency in Hertz (Hz). The pitch (F0) contours show the four tonal variations: high (black continuous), fall (black dashed), low (white continuous), rise (white dashed). Two dotted light grey auxiliary lines are 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

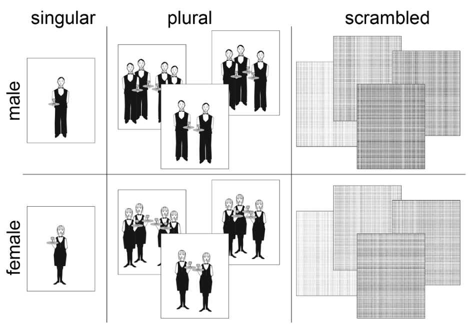


Fig. 9. Example of a set of picture stimuli used in the experiment.

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.

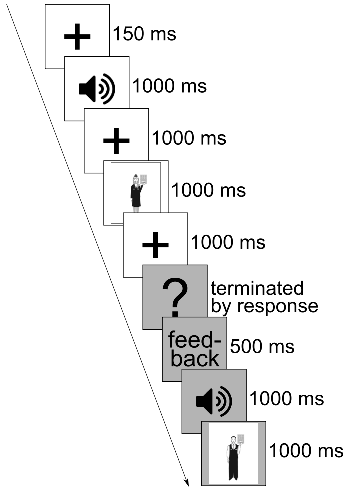


Fig. 10. Experimental procedure. Simple learning and control trials consisted of a sound stimulus followed by the corresponding picture stimulus (white slides). Question trials had an additional question slide, subsequent feedback and a repetition of the sound followed by the correct picture (grey slides).

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**	**ɛ̝**	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		j	j
uː	uː	ʊ	ʊ	x		ɣ	
	ʉː		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.

Paper IV

Native language experience narrowly shapes pre-attentive foreign tone processing and governs rapid memory trace build-up: An ERP study

Abstract

Language experience, particularly from our native language (L1), shapes our perception of other languages around us. The present study examined how L1 experience with form and functionality mould the initial processing of foreign (L2) tone during acquisition. We investigated, in particular whether learners were able to rapidly forge memory traces for novel tonal words, visible in a preconscious word recognition effect at ~50 ms. We thus recorded learners' ERP responses while they acquired novel tonal words. We manipulated the degree of L1-L2 familiarity for the tones. Our results indicate that a rapid, pre-conscious memory trace build-up for tone is possible only at particularly high levels of functional and formal L1-L2 similarity. In comparison, a late anterior negativity for the processing of the tonal words' grammatical content was apparently unaffected by native language experience but instead influenced by lexicality, pitch prominence and entrenchment. Behavioural results showed equal learning effects for all learners and tone types, regardless of the degree of L1-L2 familiarity or pitch prominence. Together, these findings suggest that early, preconscious tone processing and memory trace build-up are strongly and narrowly influenced by L1 experience while conscious processing of tonal words is largely L1-independent. Hence, facilitative effects of L1-L2 similarity are visible mainly at pre-conscious levels and do not necessarily result in more successful L2 acquisition.

1 Introduction

The present paper contributes to the study of L2 tone processing by testing beginner learners of a tone language and examining how their native language experience influences how quickly they can acquire novel tonal words. To this effect, we investigate whether or not having a tonal native language aids in the acquisition of L2 tone. In electrophysiological data, we expect possible effects of facilitation to become visible during automatic word assessment at ~50 ms, during later, more conscious linguistic processing (i.e., at N400/anterior negativity latency), or in behavioural responses pertaining to the recognition of incorrect tones. Understanding how second language learners acquire and process second language (L2) tone is important in order to reach a full and comprehensive understanding of L2 processing. While research in this area is slowly starting to emerge, the field is still heavily understudied.

1.1 Linguistic tone

Between 40 and 70% of the world's languages have word-level tone (Maddieson, 2013; Yip, 2002), i.e., a suprasegmental feature that is added onto words to make linguistic distinctions: either lexical or grammatical. In a lexical tone language, contrastive tones distinguish lexical items. The Mandarin syllable *ma*, produced with a high level tone (i.e., T1), for instance, translates to 'mother', while it means 'horse' when produced with a fall-rise pitch contour (i.e., T3). In a grammatical tone language, on the other hand, tones express grammatical features. In Somali, for instance, the change from a non-high to a high tone on the ultimate vowel in a noun translates to a change from nominative to genitive case (Banti, 1989). That is to say, tones can have different functions depending on the type of tone language in which they appear. Tone languages also differ from each other with respect to acoustic features of the tone. Register tone languages, on the one hand, make use of tones that can predominantly be distinguished according to pitch level (e.g., Yoruba: high, mid, low). Contour tone languages, on the other hand, distinguish tones according to pitch movement as well as pitch level (e.g., Cantonese: high, mid, low, mid-rise, low-rise, fall). Tone is perhaps most well-known from the Asian language families but it also plays an important role in many African and Native American languages and even in a number of European languages. While some tone languages are small or even facing extinction, others are very much alive and thriving. In fact, the language with the globally largest number of native speakers is the tonal Mandarin Chinese (>920 million native language (L1) speakers, cf. Eberhard, Simons & Fennig, 2020). It is estimated that more than half of the world's population has a tonal native language (Fromkin, 1978).

1.2 Perception of tone in second language learners

The global prominence of tone languages automatically brings along a large number of learners who acquire tone as part of a second language. For Mandarin alone, there are an estimated number of 200 million L2 speakers; and for Hausa, the largest African tone language, 25 million people are assumed to be L2 speakers (Eberhard et al., 2020). Learning a language with non-native tone is challenging. Maybe surprisingly, difficulties in this context arise not only in L2 tone production but also in its perception, processing and acquisition (e.g., mapping tones onto syllables). With respect to perception, even advanced learners of a tone language can have tone perception accuracies of below 30% for tones which native speakers can identify with nearly 100% certainty (Yang & Chan, 2010) with an average of below 70% and a large standard deviation (i.e., 47%, Pelzl, Lau, Guo & DeKeyser, 2020). Tone perception appears to be moulded by previous language experience: L2 tone types and tonal features which have functional relevance in the listeners' L1 stand out as more perceptually salient and can more easily be distinguished (cf. e.g., So & Best, 2014). As a consequence, speakers of contour tone languages relatively easily differentiate foreign tone contrasts with the help of both tone height and, importantly, tone direction cues (Gandour, 1983). Speakers of non-tonal languages, in comparison, classify tone contrasts predominantly with respect to tone height (e.g., Gandour, 1983; Huang & Johnson, 2010; Pelzl, Lau, Guo, & DeKeyser, 2019), as the tones' directionality does not have a functional word-level relevance in their L1. This often impairs non-tonal learners' in their perception of L2 tone compared to listeners from a contour tone language (e.g., Burnham, Kasisopa, Reid, Luksaneeyanawin, Lacerda, Attina, Xu Rattanasone, Schwarz & Webster, 2015; Liu, 2013; Qin & Mok, 2011; Yu, Li, Chen, Zhou, Wang, Zhang & Pi, 2019).

1.3 Perceptive processing of tone in second language learners

The L1-based difference in conscious L2 tone perception is likely due to the underlying pre-attentive processing of known tone features. Shen and Froud (2019), correspondingly, found a mismatch negativity (MMN), indicating the pre-attentive detection of a deviant tone in a series of standard tones, and a P300, showing a shift in attention upon the detection of deviant stimuli for functionally relevant tone contrasts in native speakers. For non-tonal learners and non-learners, in contrast, functionality did not seem to influence MMN amplitudes. Instead, MMNs were influenced by pitch intervals. Learners with a tonal L1, finally, showed MMNs that varied as a function of both tone functionality and pitch interval (Yu et al., 2019). This reinforces the behavioural perception results and suggests a relatively strong influence of the learners' L1 in conscious and preconscious tone perception and processing. Interestingly, the influence of L1-shaped perception is retained, albeit to a lower degree, even in relatively advanced learners (Shen & Froud, 2019).

1.4 Linguistic processing of tone in second language speakers

While some research has thus been carried on conscious foreign tone perception, as well as pre-attentive perceptive processing and the importance of the L1 in this context, very little attention has been paid to how such perceptual advantages and disadvantages, respectively, influence the linguistic processing of tonal L2 words. An interesting component in this context is a preconscious word recognition response as early as 50 ms after stimulus recognition point. This response is related to rapid memory trace formation for new words very early on in the acquisition process and has previously been seen to be influenced by L1-L2 similarity (see below). Looking at this early component thus allows for a study of automatic linguistic processing of tonal L2 words in novice learners. Secondly, more conscious linguistic processing of semantic and grammatical information is often visible at around 400 ms. Here, the established N400 component measures semantic processing while the (left) anterior negativity ([L]AN) illustrates grammatical processing. These preconscious and conscious components seem highly suitable for studying linguistic processing of tonal contrasts (see below).

1.4.1 Pre-attentive processing of tone: word recognition component

In recent years, research on rapid language learning has found an extremely early linguistically relevant component just ~50 ms after the word disambiguation point. The so-called word recognition component is sensitive to lexicosemantic (e.g., MacGregor, Pulvermüller, van Casteren & Shtyrov, 2012; Shtyrov & Lenzen, 2017) and syntactic (e.g., Herrmann, Maess, & Friederici, 2011) properties of spoken words and has been related to the automatic assessment of words' linguistic properties, such as lexicality status or syntactic category. The effect is very stable and has been observed in many different languages (e.g., English: Shtyrov & Lenzen, 2017, Finnish: Kimppa, Kujala, Leminen, Vainio & Shtyrov, 2015, German: Herrmann, Maess, Hasting, & Friederici, 2009, Danish: Partanen, Leminen, de Paoli, Bundgaard, Kingo, Krøjgaard & Shtyrov, 2017, Chinese: Yue et al., 2014), in different paradigms (e.g., in ignore conditions: Shtyrov & Lenzen, 2017, in attend conditions: Kimppa et al., 2015, with tasks: Herrmann, Maess, & Friederici, 2011, in oddball paradigms: MacGregor, Difrancesco, Pulvermüller, Shtyrov & Bohr, 2015, for single word presentations: Partanen et al., 2017, sentence presentations: Herrmann, Maess, Hahne, Schröger & Friederici, 2011 or in acquisition contexts: Gosselke Berthelsen, Horne, Shtyrov & Roll, 2020a) and for different listener populations (e.g., healthy adults: MacGregor et al., 2012, children: Partanen et al., 2017, aphasics: MacGregor et al., 2015, or L2 learners: Kimppa, Shtyrov, Hut, Hedlund, Leminen & Leminen, 2019). Changes in the word recognition component due to memory trace formation appear within minutes of language acquisition or contact. Importantly for the present study, the component also appears well-suited for studying tone processing. To this effect, Yue et al.

(2014) played Mandarin words (e.g., *peng*3, ‘to hold up in two hands’) and related pseudowords (e.g., *pang*3) to native listeners in a passive listening paradigm. They found an initially reduced negativity to scarcely presented pseudowords which quickly (i.e., within minutes) increased and turned into a larger negativity for pseudowords compared to real words. The same activation pattern is known previously from non-tonal pseudowords (e.g., Kimppa et al., 2015) and is believed to signal enhanced activation for pseudowords due to an ongoing process of auditory memory trace formation. The word recognition component has not been studied extensively for L2 learners. Yet, Gosselke Berthelsen et al. (2020a) mention increased neural activity for pseudowords – or decreased activity for meaningful novel words – in learners with a tonal native language, suggesting an impact of language experience in this early, exogenous component. However, further, more detailed research into this component for both L1 and L2 tone processing is needed.

1.4.2 Conscious processing of tone: Negativities at 400 ms

Following early, preconscious responses in language processing, conscious responses are elicited for linguistic stimuli. The N400, on the one hand, is elicited for language input with respect to meaningfulness (i.e., semantics). A strong grammatical content, on the other hand, will often evoke an anterior negativity (AN, or LAN, since often left lateralised1) instead. The N400 has larger response amplitudes for incongruent or unexpected input (Kutas & Federmeier, 2014; Kutas & Hillyard, 1980) and has been associated with facilitated lexical access (Lau, Phillips, & Poeppel, 2008). N400s have been observed in a small number of studies on L1 tone processing: In a lexical tone language, changing the tone on a target word changes its lexicosemantic content and thus turns it into a bad fit for the context. Such tone mismatches evoke N400 responses in native speakers (Brown-Schmidt & Canseco-Gonzalez, 2004; Ho, Boshra, Schmidtke, Oralova, Moro, Service & Connolly, 2019; Malins & Joanisse, 2012; Pelzl et al., 2019; Zhao, Guo, Zhou, & Shu, 2011). In a language, where tone has strong grammatical associations, like the pitch accent language Swedish, an anterior negativity (has been found for tone-suffix mismatches when there is maximal focus on the grammatical content (e.g., Söderström, Horne & Roll, 2016). No studies have yet been carried out on native speaker processing of tone mismatches in a pure grammatical tone language. Interestingly, a small number of learner studies on the processing of linguistic mismatches in different types of tones have emerged in recent years. The results are distinct from native learner processing. Even learners with good perceptual distinctions of lexical tones, for instance, show no evidence of an N400 after tone mismatches (Pelzl et al., 2019, 2020). The same is observed for beginner and intermediate learners of pitch accent languages without specific training (Gosselke Berthelsen et al., 2018; Hed, Schremm, Horne, & Roll, 2019). All of the tested L2 learners had non-tonal L1s. A further study on grammatical tone in L2 learners compared tonal and non-tonal learners of foreign tone (Gosselke Berthelsen, Horne, Shtyrov & Roll, 2020b) and found that pictures which were mismatched with the tone of a preceding word evoked an N400[2] only in learners with a tonal native language. This thus far sparse data suggests that native language experience with tones not only affects the perception and perceptive processing of tones but also their linguistic processing. For speakers of a tone language, tones are presumably stored in the left planum temporale which might be a prerequisite for the rapid processing of foreign tones (Schremm, Novén, Horne, Söderström, van Westen & Roll, 2018). With intensive tone-focused training, however, learners with a non-tonal L1 might be able to overcome native-language biases and produce conscious tone-mismatch-related ERP responses for L2 tone (Hed et al., 2019).

1.5 The present study

Briefly summarising the above, linguistic tone is an important feature in the world's languages and its perception, processing and acquisition are fairly understudied. This is true for tone languages in general, especially those with grammatical tone, but even more so for tone languages in L2 contexts and for the linguistic (i.e., not purely perceptual) processing of acquired tonal words. We know from the small number of studies that are available that L1 experience seems to shape tone perception and that this has consequences for both the preconscious processing of tone differences and the acquisition and linguistic processing of tonal words. We also know that advanced learners can to some degree overcome L1-based perception biases and that the influence of the L1 is therefore strongest in beginner learners. Studies on tone perception and perceptive processing suggest that the influence of the L1 is rather strict, such that there is facilitation only for tones that are similar in form and functionality in L1 and L2. Nevertheless, it is largely unknown how this translates to rapid word trace formation and linguistic processing of tonal words. Although results of the few existing studies cautiously point to a possible influence of language experience, it is fairly unclear how similar L1 and L2 tone need to be for a facilitated processing of tonal L2 words (e.g., differences in amplitude of N400/AN or word recognition component) to occur. The present study used ERPs to investigate the initial acquisition of tonal words and examined which degree of similarity would be necessary for facilitation effects to appear. For this purpose, we constructed tonal pseudowords which we taught to native speakers from closely related tonal and non-tonal languages (Swedish and German) over the course of two days. The tones in the present study were level tones (high and low) and contour tones (fall and rise) that were implemented on a CVC syllable (i.e., consonant-vowel-consonant). The tonal learners' native language was a contour tone language with two contrastive tones whose shape is influenced by dialectal differences, the presence or absence of focus as well as sentence intonation (Bruce, 1977). As a result, the Swedish tone system is relatively complex but strongly based on tone movements (Bruce, 2005). Therefore, if L1-L2 tone similarity is the determining factor in L2 tone acquisition, Swedish speakers should show effects of facilitated acquisition (i.e., faster word trace formation visible at ~50 ms or a decreased N400/[L]AN for acquired words) only for words with contour tones. If, however, general experience with word-level tone is sufficient to facilitate L2 tone acquisition, we should see no differences between tone types in the tonal L1 learners but still clear differences between tonal and non-tonal learners. If, finally, L1 experience does not influence L2 tone acquisition in its initial stages, there should be no processing differences between learners from different L1 backgrounds. It is, further, also possible that the early (preconscious) and late (conscious) response are affected differentially by L1 background.

2 Methods

2.1 Participants

Forty-eight healthy, right-handed adults (25 female, mean age 23.7) with normal or corrected-to normal vision and normal hearing (defined as pure-tone hearing thresholds ≤20 dB Hearing Level (ISO, 2004)) were recruited for the study. All tests were carried out at the Lund University Humanities Lab and most participants were students at Lund University. Half of the participants had a tonal L1, Swedish, the other half a non-tonal L1, German. The tonal and non-tonal participants were each divided into two learner groups: high/fall learners (i.e., learners who acquired target words with high and falling tones, where low and rising tones served as controls) and low/rise learners (i.e., learners who acquired target words with low and rising tones, where high and falling tones served as controls). All four groups (i.e. tonal L1 high/fall, tonal L1 low/rise, non-tonal high/fall, non-tonal low/rise) were matched for age, gender, socioeconomic status (Hollingshead, 1975), working memory span (Unsworth, Heitz, Schrock, & Engle, 2005), their perception of non-tonal phonological contrasts (vowel duration, mean accuracy 97.1%) and their discrimination of extra-linguistic pitch (piano tones, mean accuracy 92.5%). All subjects were remunerated for their participation. One tonal L1 high/fall participant reported previous exposure to a foreign tone language and one participant from the non-tonal L1 low/rise group chose to discontinue the experiment. The data from both participants was consequentially excluded.

2.2 Stimuli

To test tone processing and tone word acquisition, we created 24 tonal pseudowords. The words had a simple CVC structure (consonant-vowel-consonant), e.g., /siːs/, and were constructed by equalising initial consonants, vowels and final consonants for length and loudness and splicing them together in Praat (Boersma, 2001) with 10 ms transition phases. All individual sounds were recorded in an anechoic chamber by a male speaker of Russian. Consonants were preceded or followed by the same two dummy vowels. Russian phonemes were chosen to avoid differential carry-over effects which German or Swedish phonemes could have evoked. All test words were perceived equally well and correctly classified as pseudowords by eight speakers of German and three speakers of Swedish. In a final step, we used pitch-manipulations to add two level tones (high: 138 Hz and low: 98 Hz) and two contour tones, a rise (98 Hz to 138 Hz) and a fall (138 Hz to 98 Hz). Onset of pitch movement was aligned with the onset of the vowel (cf. Figure 1).

In order to for the stimuli to contain linguistic content, we taught them through association with meaningful pictures. The pictures we constructed for this purpose showed people in 24 different professions. For each profession, versions of the pictures illustrating gender (i.e., masculine = male worker, feminine = female worker) and number (i.e., singular = one worker, plural = two, three or four workers) were created. In addition to the profession pictures, diffuse gray pictures were added as meaningless control pictures. See Figure 1.

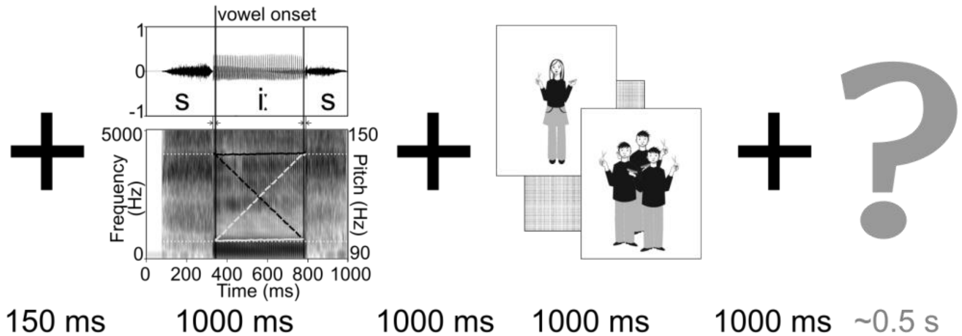


Figure 1. Example of an auditory stimulus and a subset of visual stimuli embedded in the experiment procedure. The question mark in gray simplistically illustrates the addition of a question to 12% of all trials. In question trials, a question concerning the correctness of the previous word-picture pair was followed by overt feedback and a repetition of the auditory stimulus and the associated visual stimulus.

2.3 Procedure

Each artificial L2 learner participated and learned twenty-four words on two consecutive days. At their arrival, participants gave written consent about participating in the study. During EEG application on the first day, they filled in a number of background questionnaires. They were subsequently seated at a fixed distance from a computer screen on which the visual stimuli, fixation crosses, questions and feedback where presented. These were created in and controlled by E-Prime 2 stimulation software (Psychology Software Tools Inc., Sharpsburg, PA). All subjects were asked to keep their index fingers on a response box on a table in front of them to answer questions, when prompted. The auditory stimuli in the experiment 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 Kjær 2231 sound level meter with a 4134 microphone in a 4153 Artificial Ear. At the very beginning of the first learning session, participants received explicit instructions regarding the learning paradigm and conducted a test phase which illustrated the learning procedure with the help of Spanish words with gender and number suffixes (i.e., *mechanico/a(s)*, mechanic.MAS/FEM(PL), 'mechanic(s)' and *arquitecto/a(s)*, architect.MAS/FEM(PL), 'architect(s)'). Afterwards, the learning procedure commenced with an auditory stimulus followed by a meaning-giving picture (cf. Figure 1). Stimulus onset asynchrony was 4.15 seconds. Meaning assignment was strongly regulated so that the initial consonant always assigned lexicosemantic information (i.e., the profession), while the vowel and the tone were associated with the grammatical categories gender and number. Across participants, we counterbalanced the distribution of professions as well as whether vowel or number was associated with gender or tone. Two vowels and tones were part of the learning paradigm (target words) while the other two were used in control words which were presented with scrambled pictures. Control words served as comparison stimuli to distinguish learning from familiarisation. For vowels, a/ɛ or i/u were paired together. Each pair was equally often part of the target or control words. Tones were split up in the same way, ensuring that word onsets for the target words for each participant had identical pitch levels. This resulted in an high/fall group and a low/rise group. In the high/fall group, target words had high or falling pitch, while the controls had low and rising tones. For the low/rise group, the pattern was inversed. Finally, to elicit behavioural measures to assess the learning progress, in approximately six percent of all trials, the stimulus and the picture were mismatched. Together with six percent of the congruous trials, these served as question trials. In a question trial, participants were asked to judge the correctness of the previous word-picture pair. When participants answered with help of the button box, overt feedback was provided. Depending on response times, question trials were approximately five seconds longer than non-question trials. On each day, every stimulus (24 targets and controls) was repeated 30 times in non-question trials and at least once in a question trial (30 in total). Participants were offered a longer break after every 10 repetitions (~40 minutes).

2.4 Electrophysiology

During the acquisition paradigm, the participants' brain activity was recorded with the help of 64 Ag-AgCl EEG electrodes mounted in an electrode cap (EASYCAP GmbH, Herrsching, Germany), a SynAmps2 EEG amplifier (Compumedics Neuroscan, Victoria, Australia), and Curry Neuroimaging Suite 7 software (Compumedics Neuroscan). We monitored eye movements with horizontal and vertical bipolar electrooculogram electrodes (EOG). Impedance at scalp channels was kept below 3 kΩ and below 10 kΩ for the eye channels. Online reference was left mastoid (M1) and frontocentral electrode AFz served as ground. We recorded EEG with a 500 Hz sampling rate using DC mode and an online anti-aliasing low pass filter at 200 Hz. We subsequently filtered the data with a 0.01 Hz high pass and a 30 Hz low pass filter. ERP epochs were extracted at word disambiguation point (vowel onset) with a length of 1200 ms including a 200-ms baseline. An independent component analysis (ICA; Jung, Makeig, Humphries, Lee, McKeown, Iragui & Sejnowski, 2000) was conducted and components related to eye artefacts and bad channels were removed. Epochs still exceeding ± 100 µV were discarded. Due to the high number of trials in the paradigm (i.e., 3156 per participant), the ICA components were well defined. The rejection of, on average, 5 out of 66 components was sufficient to correct the data to such a degree that only few trials (M = 35 per participant) exceeded the 100 µV threshold.

2.5 Statistical analysis

2.5.1 Behavioural data

To test for possible effects of L1-L2 similarity in the behavioural results, we analysed the behavioural responses to tone mismatch trials focusing on differences between words with contour tones and words with level tones. To this effect, we separately submitted mean data for the behavioural variables 'Response Accuracy' and 'Response Times' to two mixed analyses of variance (ANOVA) with the experimental factor 'Tone Type' (contour tones vs. level tones), the temporal factor 'Day' (day 1 vs. day 2) and the between-subject factors 'Learner Group' (tonal L1s vs. non-tonal L1s) and 'Target Tone Group' (high/fall vs. low/rise). For Response Times, only responses to correctly identified mismatches were considered. Response Times were normalised through log transformations and for accuracy, d′ scores were computed.

As two participants were excluded from the study, the Target Tone subgroups were of different size (11 or 12). We used mean imputation wherever same sized target groups were necessary for the statistical analysis.

2.5.2 ERPs

For the ERP data, we selected two time windows (50-70 ms and 400-600 ms) where we expected to observe influences of L1-L2 similarity on tonal or tone features on word processing, based on previous literature. Using these pre-defined time-windows, we conducted cluster-based permutation tests for the factor 'Tone Type'. We submitted mean ERP amplitudes of block 1 (on day 1) from both participant groups, together and separately, for the selected time windows and conditions (i.e. words with contour tones compared to words with level tones) to a permutation analysis using the nonparametric cluster-based permutation approach implemented in Fieldtrip (Maris & Oostenveld, 2007). We ran 1000 random permutations of the data with the Monte Carlo method to account for large data sets and considered clusters of three or more electrodes with a p-value of < 0.05 significant. We additionally tested for interactions with 'Learning' (target word vs. control word) in the permutation analysis in order to see whether target words differed from control words. The interaction was particularly important for the word recognition effect at 50 ms, where tonal learners have earlier been seen to automatically dissociate target words from control words (Gosselke Berthelsen et al., 2020a).

If significant clusters emerged in any of the analyses, we carried out mixed ANOVAs to test for possible interactions with temporal and between-subject factors. Thus, we computed mean ERP amplitudes for the cluster electrodes in the relevant groups and submitted them to a mixed ANOVA with the experimental within-subject factors 'Tone Type' and 'Learning', the temporal factors 'Day' and 'Block', as well as the between-subject factors 'Learner Group' (if applicable) and 'Tone Target Group'. Greenhouse-Geisser correction was used when applicable. Main effects and interactions were considered significant at a p-value of < 0.05. For pairwise comparisons, False Discovery Rate (FDR) corrections (Benjamini & Hochberg, 1995) were applied.

3 Results

3.1 Behavioural results

For Response Accuracy, there was no significant main effect of the factor Tone Type and no significant interactions with the factors Tone Type, Learner Group or Tone Target Group. Main effect of the temporal factor Day, $F(1,44) = 44.18$, $p < .001$, showed evidence of learning regardless of which sets of tones the learners acquired or which tone types were tested. Bonferroni-corrected pairwise comparisons revealed a significant improvement in Response Accuracy from day 1 (M = 60.5%, SD = 25.4) to day 2 (M = 68.8%, SD = 29.6).

For Response Times, a significant interaction of Tone Type and Tone Target Group, $F(1,44) = 4.63$, $p = .037$, broke down into a main effect of Tone Type in the low/rise group, $F(1,22) = 6.69$, $p = .015$. Mismatches with the pictorial referent based on low tones (M = 1515 ms, SD = 986) were significantly faster detected than errors based on rising tones (M = 1722 ms, SD = 1093). There was no main effect of the factor Tone Type and no significant interactions with the factors Tone Type, Learner Group or Tone Target Group. Main effect of the temporal factor Day, $F(1,44) = 35.53$, $p < .001$, showed evidence of learning regardless of which tones types the learners were taught or tested on. To this end, there was a significant improvement in Response Times from day 1 (M = 2056 ms, SD = 329) to day 2 (M = 1427 ms, SD = 153). As mentioned above, only Response Time values for accurately detected tone errors were analysed.

3.2 ERP results

3.2.1 50-70 ms

For the early time window, an interaction between Tone Type and Learning, i.e. comparing level and contour tones in target and control words produced a significant central electrode cluster (FC2, FC4, C1, Cz, C2, C4, CP1, Cpz, CP2), $p = .022$ in the tonal L1 group. See ERPs and topographies for the interaction in Figure 2. No comparable cluster was found in the non-tonal L1 group or for all participants, collectively. The permutation analysis did not produce any significant clusters for differences between level and contour tones without an interaction with Learning, neither for all participants collectively nor for either of the participant groups.

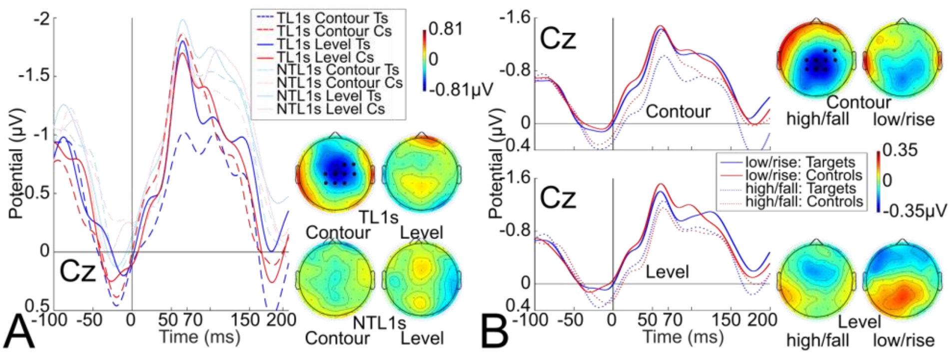


Figure 2. ERPs and subtraction topographies (controls minus targets) for the 50-70 ms effect.
A. Session initial ERPs at central electrode Cz and topographies for the Tone Type by Learning interaction cluster of the permutation analysis (electrodes marked in black) in the tonal L1 (TL1) group (TL1; top) and comparable topographies for the non-tonal L1 group (NTL1; bottom, N.S.). B. Global ERPs and topographies for the early effect in the tonal L1 group by Tone Type: Contour (top) and Level (bottom). (Ts = Targets, Cs = Controls).

Further investigating the significant cluster of the tonal L1 group in a mixed ANOVA, a Tone Type * Learning * Target Tone Group interaction suggested that the observed effect for contour tones (i.e., controls were more negative than targets) was significant in the high/fall group only. Secondly, an interaction with time indicated that only amplitudes of control words changed over time, turning less negative. For detailed results, see Table 1.

Table 1.- All significant results of the mixed Analysis of Variance (ANOVA) analysis for the tonal L1 learners' ERPs in the early time window (TTG = Target Tone Group, H/F = high/fall learners)

MAIN EFFECTS, INTERACTIONS & PAIRWISE COMPARISONS	**F**	**p**	**MEANS IN µV**			
Tone Type * Learning	9.99	.005				
Contour Tones: Learning	17.38	<.001	Control	-1.12	Target	-0.95
Level Tones: Learning	122.58	<.001	Control	-1.04	Target	-1.04
Tone Type * Learning * TTG	7.56	.012				
H/F: Tone Type * Learning	12.58	.005				
H/F, Contour Tones: Learning	15.44	.002	Control	-1.03	Target	-0.78
Learning	6.83	.016	Control	-1.08	Target	-0.99
Block	3.20	.021				
Pairwise comp.: Block 1 vs Block 4		.008	Block 1	-1.17	Block 4	-0.92
Learning * Block	4.33	.004				
Control: Block	5.45	.001				
Pairwise comp.: Block 1 vs Block 3		.012	Block 1	-1.13	Block 3	-1.03
Block 1 vs Block 4		.006	Block 1	-1.13	Block 4	-0.97
Block 1 vs Block 6		.011	Block 1	-1.13	Block 6	-0.90
Block 5 vs Block 6		.008	Block 5	-1.18	Block 6	-0.90

3.2.2 400-600 ms

For the second time window, permutation analysis produced two significant clusters for the comparison of words with contour tones and words with level tones for the first 20 minutes of the first session in all participants. There was a significant frontocentral cluster (AF3, AF4, AF8, F5, F3, F1, Fz, F2, F4, FC5, FC3, FC1, FC2, FC4, FC6, C3, C1, C2, C4), $p < .001$, as well as a significant posterior cluster (FT7, FT8, TP7, CP6, TP8, P7, P5, P8, PO7, POz, PO8, Oz), $p < .001$. See Figure 3.

A mixed ANOVA of the mean ERP amplitudes of the frontocentral cluster for all participants in 20-minute blocks yielded a main effect and interactions for Tone Type, for Learning, and for the temporal factors Block and Day, cf. Table 2. With regards to Tone Type, level tones turned out to be more negative than contour tones. This difference was stronger in the tonal L1 group than in the non-tonal L1 group. For Learning, we found that target words were overall more negative than control words. The difference was again stronger in the tonal L1 group and also stronger on day 2 than on day 1. Finally, a general decrease of the negativity of the effect was observed over blocks and days.

Table 2. All significant results of the mixed Analysis of Variance (ANOVA) analysis for ERPS of the frontal cluster in the late time window. (LG = Learner Group, TL1 = tonal L1 learners; NTL1 = non-tonal L1 learners, TTG = Target Tone Group, H/F = high/fall learners, L/R = low/fall learners, D = Day, B = Block, TT = Tone Type).

MAIN EFFECTS, INTERACTIONS & PAIRWISE COMPARISONS	F	p	MEANS IN µV			
Tone Type	191.83	<.001	Level	-3.24	Contour	-2.86
Tone Type * Learner Group	4.23	.046				
TL1: Tone Type	204.65	<.001	Level	-3.27	Contour	-2.74
NTL1: Tone Type	50.34	<.001	Level	-3.25	Contour	-2.86
Tone Type * Block * Learner Group	3.01	.015				
Block 4: Tone Type * Learner Group	5.23	.027				
Block 4, TL1: Tone Type	54.76	<.001	Level	-3.33	Contour	-2.75
Block 4, NTL1: Tone Type	19.76	<.001	Level	-3.27	Contour	-2.94
Tone Type * Day * Block * LG * TTG	3.09	.016				
H/F: Tone Type * Day * Block * LG	2.73	.036				
H/F, D1: Tone Type * Block * LG	3.67	.010				
H/F, D1, B2: Tone Type * LG	5.19	.033				
H/F, D1, B2, NTL1: TT	25.62	.001	Level	-3.74	Contour	-3.02
H/F, D1, B5: Tone Type * LG	7.99	.010				
H/F, D1, B2, TL1: TT	50.52	<.001	Level	-3.82	Contour	-3.08
H/F, D2: Tone Type * Block * LG	2.51	.045				
H/F, D2, B4: Tone Type * LG	6.34	.020				
H/F, D2, B4, TL1: TT	20.44	.001	Level	-3.34	Contour	-2.62
Learning	36.73	<.001	Control	-2.89	Target	-3-17
Learning * Learner Group	4.37	.011				
TL1: Learning	33.61	<.001	Control	-2.81	Target	-3.20
NTL1: Learning	7.79	.011	Control	-2.96	Target	-3.15
Learning * Day	7.87	.007				
Day 1: Learning	20.55	<.001	Control	-3.16	Target	-3.38
Day 2: Learning	38.55	<.001	Control	-2.61	Target	-2.97
Learning * Block * Learner Group	2.87	.023				
NTL1: Learning * Block	3.83	.006				
NTL1, B2: Learning	5.21	.032	Control	-2.88	Target	-3.09
NTL1, B4: Learning	12.09	.002	Control	-2.96	Target	-3.25
NTL1, B5: Learning	6.77	.016	Control	-2.87	Target	-3.09
NTL1, B6: Learning	12.35	.002	Control	-2.92	Target	-3.24
Tone Type * Learning * Day * TTG	5.94	.019				
L/R: Tone Type * Learning * Day	12.65	.002				
L/R, Level: Learning * Day	14.53	.001				
L/R, Level, D1: Learning	5.49	.029	Control	-3.10	Target	-3.33
L/R, Level, D2: Learning	23.97	<.001	Control	-2.69	Target	-3.22
Block	3.47	.013				
Multiple comparisons		.026	Block 1	-3.16	Block 5	-2.92
Day	23.73	<.001	Day 1	-3.27	Day 2	-2.79
Day * Target Tone Group	5.23	.027				
H/F: Day	29.11	<.001	Day 1	-3.54	Day 2	-2.83
Day * Block * Learner Group * TTG	3.30	.022				
TL1: Day * Block * TTG	4.19	.011				
TL1, H/F: Day * Block	3.20	.049				
TL1, H/F, D2: Block	5.17	.021				
Pairwise comparisons		.026	Block 1	-3.26	Block 2	-2.90
		.004	Block 1	-3.26	Block 3	-2.56
		.016	Block 1	-3.26	Block 5	-2.44
		.013	Block 1	-3.26	Block 6	-2.43
		.037	Block 2	-2.90	Block 6	-2.43
		.013	Block 3	-2.56	Block 4	-2.98

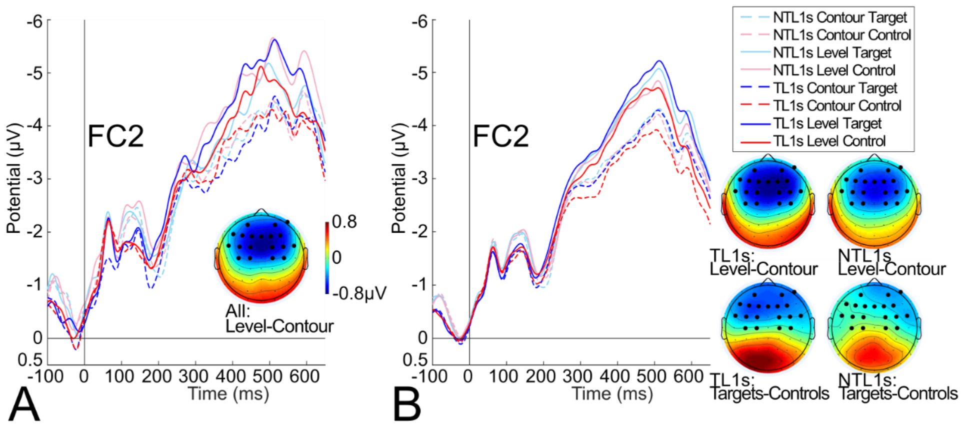


Figure 3. ERPs and subtraction topographies for the anterior negativity at 400-600 ms.
A. Session initial ERPs at frontocentral electrode FC2 and topography (level−contour) for the significant cluster of the permutation analysis (electrodes marked in black) found for all participants. B. Global ERPs and topographies for the Tone Type effect (level−contour; top) and the Learning effect (target−control; bottom) for all trials in the tonal (TL1) and non-tonal L1 (NTL1) group.

For the posterior cluster, the effects were virtually indistinguishable from those of the frontocentral cluster but reversed in polarity. All of the main effects were identical as were the crucial interaction clusters. Only two unique interactions emerged in the posterior cluster. We therefore chose to treat the positivity as a dipole effect and, for the sake of brevity, present the observed significant effects and interactions for the posterior positive cluster as supplementary material instead of in the main text.

4 Discussion

4.1 Word recognition component: transfer effects

There was a clear effect of native language experience in the pre-attentive word assessment component at ~50 ms. The facilitation effect was very narrow, possibly more so than we originally anticipated, and as such only found for the highest degree of L1-L2 similarity. Our tonal learners did not show indications of facilitated word acquisition for all tonal target words or for all target words with moving contour tones, but instead only for target words with a falling tone. This became apparent in a reduced negativity for target words with falling tones. We assume that this reflects a successful, ultra rapid word trace formation for target words with falling tones such that they become processed real-word-like already within the first twenty minutes of acquisition, hence the decreased amplitude. A similar decrease has been found for real words in Mandarin speakers (Yue et al., 2014). After four minutes of word and legal pseudoword repetition, neural activity to real words appeared to become reduced3. For non-tonal words, the same effect is reported in Kimppa et al. (2015). We therefore interpret the decreased effect size for narrowly L1-facilitated target words in the present study as evidence that these words are acquired and processed like real words, extremely fast. Pseudowords and non-facilitated target words, on the other hand, could not be acquired equally rapidly and therefore evoked an increased negativity. This negativity likely signals an ongoing, incomplete memory trace formation process for untaught and non-facilitated words. The same increased activity for pseudowords has previously been attested where pseudowords and real words were repeatedly presented both in the context of tonal (Yue et al., 2014) and non-tonal words (Kimppa et al., 2015).

Interestingly, we found a facilitation effect only for words with falling tones in the tonal learners. In accordance with studies on the importance of L1 tone shape (and functionality) in tone perception (e.g., Burnham et al., 2015; Huang & Johnson, 2010), this strongly suggests the particular importance of falling tones in the learners' native language, Swedish. While it has previously been stipulated that learners of a contour tone language have facilitated, perceptual access to contour tones on the whole (e.g., Gandour, 1983), this concept of general facilitation is challenged by the current data. It seems as though L1 experience shapes L2 tone acquisition very narrowly, at least with respect to preconscious linguistic processing and memory trace formation. Only tonal L2 words with shapes that have direct functional relevance in the L1 make ultra rapid acquisition and word trace formation possible. Although Swedish word accents have complex, combinatory rise-fall patterns in many dialects (Bruce & Gårding, 1978; Gårding & Lindblad, 1973), influenced by focus (Bruce, 1977), Bruce (1977, 1987, 2005) argues that Swedish word accents are characterised as falls (or high-low pitch gestures) with different timings with respect to the stressed vowel (H+L* vs. H*+L). The present data supports this important minimal distinction, showing that language experience prepares Swedish listeners to preconsciously respond to and acquire the functionality of falling pitch but not that of rising or level pitch. Furthermore, the data suggests that a L1 facilitation of L2 tone processing at the preconscious level is very closely tied to functional acoustic patterns. Thus, facilitation based on L1-L2 transfer does occur for pre-attentive L2 tone processing, instantaneously in fact, but only at the highest level of formal and functional similarity.

4.2 Anterior negativity and response times & accuracy: no transfer effects

While we found clear influences of language experience on the pre-attentive processing of foreign tone, no such influences were visible in later, conscious processing or behavioural responses. All four learner subgroups, regardless of native language background or type of target tones, were equally accurate and quick at detecting tone mismatches. Accuracy levels transcended chance levels but remained relatively low (>70%) which attests to the general difficulty of L2 tone acquisition (cf., e.g., Yang & Chan, 2010). Learners found it difficult to make use of the rules underlying the tones' grammatical meaning offline. They could, however, use the tones online to differentiate between target and control words, visible in an increased anterior negativity (AN) to target words. We find an anterior negativity rather than an N400 due to the words' prominent grammatical content which was specifically targeted in the experimental task. While tones are an important cue in target and control word distinction, vowels also differentiate meaningful and meaningless words. The AN is, therefore, not necessarily based on tone processing alone but is likely also influenced by the vowels. We find lower AN amplitudes for untaught pseudowords than for taught, meaningful novel words, due to the fact that only meaningful words carry grammatical content that needs to be processed. The increase in amplitude is thus very probably based on the processing of grammatical content. It is likely that the amplitude change occurs due to the increased difficulty of processing and integrating the rules of the novel grammar compared to the processing of morphosyntactically and semantically empty control pseudowords. This assumption is corroborated by the fact that grammar errors, which naturally are even more effortful to process than canonical grammar, evoke a significantly larger anterior negativity (e.g., Molinaro, Barber, Caffarara & Carreiras, 2014).

The amplitudes of the anterior negativity were not only influenced by target and control word status but also, maybe more importantly and certainly more strongly, by tone type: the negativity was reduced for words with contour tones compared to words with level tones. This was the case in both learner groups irrespective of whether words were targets or controls. We suggest that contour tones, at least in the context of the present study, are more perceptually prominent than level tones, regardless of whether they are assigned meaning or not. This results in eased processing and a reduced negativity. In the learning paradigm in our study, all target words had the same pitch onset and could be differentiated based on whether the pitch stayed at onset level (level tone) or started moving (contour tone). The same was true for control words. That is to say, words or tones could be distinguished at movement onset, which supposedly made contour tones more perceptually prominent or salient than level tones. This factor, maybe study-specific, maybe general, likely reduced the processing load of contour tones and facilitated the rule-based integration of the tone's grammatical content. This becomes visible in a decreased AN. Similar facilitation effects for acoustic properties of tones, outside of the context of grammar, have previously been reported in the comparison of high and low pitch (piano tones, Gosselke Berthelsen et al., 2018) and of tones with steep (high-low) and moderate (mid-low) falls (Swedish pitch accents, Gosselke Berthelsen et al., 2018; Kochančikaitė, Shtyrov & Roll, 2020). The high tones or high onsets in these studies resulted in reduced negativities, suggesting a relative stronger salience of high pitch and/or steeper movement and thus less effortful processing. Hence, at least for the studied listener groups in these as well as the present study (i.e., Swedes and Germans), certain general characteristics of tones appear to shape the tones' perceptual prominence and as a result their processability. Interestingly, but maybe not surprisingly, the pitch salience-related processing load seems to add on to the processing that is prevalent for the stimuli it is associated with. In the present study, the words bear salient grammatical meaning which results in a grammar-based anterior distribution of the negativity. The same reduced anterior negativity for perceptually prominent tones has been observed for the processing of natural language with a focus on grammar associations (e.g., Gosselke Berthelsen et al., 2018). In grammar-devoid contexts, however, or for L2 learners that do not yet have a good grasp of L2 grammar, the effect is centrally distributed (Gosselke Berthelsen et al., 2018). In strongly lexical semantic contexts, we could anticipate a more posterior distribution of the effect, as it adds onto the N400.

Finally, temporal factors also influenced AN amplitude in the present study such that amplitudes are higher on the first day of acquisition compared to the second day. The decrease in negativity is likely related to the relative processing difficulty and increased ease of integration over time and has previously been observed in association learning contexts: Bermúdez-Margaretto, Beltrán, Cuetos & Domínguez (2018) found a decreased N400 due to repetition in semantic-associative training. In accordance with their findings, all novel words were easier to process on the second day compared to the first day of the experiment due to entrenchment. This resulted in differential anterior negativity amplitudes. This effect was only significant in learners who had high/fall target words, possibly due to a general trend for overnight consolidation effects reinforced through strong transfer-facilitated consolidation in the high/fall group of the tonal L1 participants. Further between-group differences and interactions in the AN time-window were primarily expressions of differentially strong effects between learner groups. Effects tended to be stronger in tonal L1 learners, although the effects were still clearly significant in the non-tonal learners (often $p < .001$). While previous studies on intermediate L2 tone learners in natural acquisition contexts have not found evidence of linguistic processing at AN/N400 latency (e.g., Pelzl et al., 2019, 2020; Gosselke Berthelsen et al., 2018), the present study finds such effects, presumably because it examines taught and untaught words in ongoing acquisition rather than error processing. Error processing is to a stronger degree dependent on prediction, a feat that learners are often incapable of accomplishing (cf., e.g., Guillelmon & Grosjean, 2001).

Interestingly, there were no important interactions of the two main factors that influence AN amplitudes in the present study: tone type and word status. That is, they differed independently of each other. All target words had a larger AN than all control words but for both target and control words, contour tones had a stronger negativity than level tones. This shows that the ERP effect in this time window is influenced separately by both the presence of grammar and by tone prominence. Interestingly, temporal factors also influence processing at this latency such that amplitudes decreased over time. Together, this suggests that the reduced negativity for tonal words in the present data is an indicator of eased processing: semantically and morphologically empty pseudowords are easier to process than newly taught words with complex morphosyntax, prominent contour tones which are strong cues to word distinctions are easier to process than level tones, and processing of the novel words becomes easier, less resource-heavy, over time. Importantly, the processing load and integration difficulty of words with L2 tone is not contingent on experience with word-level tone. Instead, word status, pitch prominence and entrenchment appear to be universal factors that impact the conscious processing of (novel) tonal words.

5 Conclusion

The current study investigated how native language experience shapes L2 tone processing and acquisition. We found a narrow effect of L1-L2 similarity in an early word recognition component at ~50 ms, such that only tones that were identical to tonal learners' native tones in functionality and pitch shape (here, falls) were acquired ultra rapidly, i.e., within 20 minutes, and consequentially stood out against all other types of tones. For non-tonal learners, there was no difference between tone types in the early, preconscious ERPs, highlighting that differential preconscious processing and memory trace formation cannot occur without functional L1-L2 transfer. Later, conscious processing, on the other hand, was not affected by language experience. Tonal and non-tonal learners alike produced differential anterior negativities influenced by seemingly universal factors such as lexicality (novel words have larger negativity than pseudowords), pitch prominence (level tones have greater negativity than contour tones) and entrenchment (day 1 has larger negativity than day 2). The learners' late, conscious processing, manifested here in anterior negativities, was not affected by L1 experience. The same was true for the learners' behavioural responses to tone mismatches. Both groups identified tone mismatches equally rapidly and equally well. The present results highlight that preconscious processing during L2 tone acquisition can be facilitated by language experience. At the same time, they relativise the importance of transfer effects by showing that only highly similar features can lead to facilitation of preconscious processing. They further suggest that a lack of L1-L2 similarity can be overcome by later, conscious processing. This is likely the reason why transfer effects do not always appear in studies on L2 acquisition and processing. However, it is possible that the impact of preconscious processing is stronger in natural second language acquisition and that learners, depending on testing conditions, can perform better for facilitated tones or experience faster learning progression in such settings.

References

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

Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society. Series B (Methodological), 57*(1), 289-300. doi:10.2307/2346101

Bermúdez-Margaretto, B., Beltrán, D., Cuetos, F., & Domínguez, A. (2018). Brain signatures of new (pseudo-) words: visual repetition in associative and non-associative contexts. *Frontiers of Human Neuroscience, 12*(354). doi:10.3389/fnhum.2018.00354/full

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

Brown-Schmidt, S., & Canseco-Gonzalez, E. (2004). Who do you love, your mother or your horse? An event-related brain potential analysis of tone processing in Mandarin Chinese. *Journal of Psycholinguistic Research, 33*(2), 103-135. doi:10.1023/B:JOPR.0000017223.98667.10

Bruce, G. (1977). *Swedish word accents in sentence perspective.* Lund: Gleerup.

Bruce, G. (1987). How floating is focal accent? In K. Gregersen, & H. Basbøll (Eds.), *Nordic prosody IV: papers from a symposium* (pp. 41-49). Odense: Odense University Press.

Bruce, G. (1987). How floating is focal accent? In K. Gregersen, & H. Basbøll (Eds.), *Nordic prosody IV: papers from a symposium* (pp. 41-49). (Odense University studies in linguistics; Vol. 7). Odense: Odense University Press.

Bruce, G. (2005). Intonational prominence in varieties of Swedish revisited. In S.-A. Jun (Ed.), *Prosodic typology: The phonology of intonation and phrasing.* doi:10.1093/acprof:oso/9780199249633.003.0015

Bruce, G., & Gårding, E. (1978). *A prosodic typology for Swedish dialects.* (E. Gårding, G. Bruce, & R. Bannert, Eds.) Lund: Department of Linguistics, Lund University.

Burnham, D., Kasisopa, B., Reid, A., Luksaneeyanawin, S., Lacerda, F., Attina, V., . . . Webster, D. (2015). Universality and language-specific experience in the perception of lexical tone and pitch. *Applied Psycholinguistics, 36*, 1459-1491. doi:10.1017/S0142716414000496

Eberhard, D. M., Simons, G. F., & Fennig, C. D. (Eds.). (2020). *Ethnologue: Languages of the world* (23rd ed.). Dallas, Texas: SIL International. Online version: http://www.ethonologue.com.

Fromkin, V. (Ed.). (1978). *Tone: a linguistic survey.* New York, NY: Academic Press.

Gandour, J. (1983). Tone perception in Far Eastern langauges. *Journal of Phonetics, 11,* 149-175. doi:10.1016/S0095-4470(19)30813-7

Gårding, E., & Lindblad. (1973). Constancy and variation in Swedish word accent patterns. *Working Papers* (Phonetics Laboratory, Lund University)*, 7,* 36-110.

Gosselke Berthelsen, S., Horne, M., Brännström, K. J., Shtyrov, Y., & Roll, M. (2018). Neural processing of morphosyntactic tonal cues in second-language learners. *Journal of Neurolinguistics, 45*, 60-78. doi:10.1016/j.jneuroling.2017.09.001

Gosselke Berthelsen, S., Horne, M., Shtyrov, Y., & Roll, M. (2020a). Different neural mechanisms for rapid acquisition of words with grammatical tone in learners from tonal and non-tonal backgrounds: ERP evidence. *Brain Research, 1729*(146614). doi:10.1016/j.brainres.2019.146614

Gosselke Berthelsen, S., Horne, M., Shtyrov, Y., & Roll, M. (2020b). *Phonological transfer effects in novice learners: A learner's brain detects grammar errors only if the language sounds familiar.* Manuscript submitted.

Guillelmon, D., & Grosjean, F. (2001). The gender marking effect in spoken word recognition: The case of bilinguals. *Memory & Cognition, 29*(3), 503-511. doi:10.3758/bf03196401

Hed, A., Schremm, A., Horne, M., & Roll, M. (2019). Neural correlates of second language acquisition of tone-grammar associations. *The Mental Lexicon, 14*(1), 98-123. doi:10.1075/ml.17018.hed

Herrmann, B., Maess, B., & Friederici, A. D. (2011). Violation of syntax and prosody - Disentangling their contributions to the early left anterior negativity. *Neuroscience Letters, 490*, 116-120. doi:10.1016/j.neulet.2010.12.039

Herrmann, B., Maess, B., Hahne, A., Schröger, E., & Friederici, A. D. (2011). Syntactic and auditory spatial processing in the human temporal cortex: An MEG study. *NeuroImage, 57*, 624-633. doi:10.1016/j.neuroimage.2011.04.034

Herrmann, B., Maess, B., Hasting, A. S., & Friederici, A. D. (2009). Localization of the syntactic mismatch negativity in the temporal cortex: An MEG study. *NeuroImage, 48*, 590-600. doi:10.1016/j.neuroimage.2009.06.082

Ho, A., Boshra, R., Schmidtke, D., Oralova, G., Moro, A. L., Service, E., & Connolly, J. F. (2019). Electrophysiological evidence for the integral nature of tone in Mandarin spoken word recognition. *Neurophysiologia, 131*, 325-332. doi:10.1016/j.neuropsychologia.2019.05.031

Hollingshead, A. B. (1975). *Four factor index of social status.* Unpublished manuscript, Yale University, New Haven, CT.

Huang, T., & Johnson, K. (2010). Language specificity in speech perception: Perception of Mandarin tones by native and nonnative listeners. *Phonetica*, *67*, 243-267. doi:10.1159/000327392

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.*

Jung, T.-P., Makeig, S., Humphries, C., Lee, T.-W., McKeown, M. J., Iragui, V., & Sejnowski, T. J. (2000). Removing electroencephalographic artifacts by blind source separation. *Psychophysiology*, *37*, 163-178. doi:10.1111/1469-8986.3720163

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. doi:10.1016/j.neuroimage.2015.05.098

Kimppa, L., Shtyrov, Y., Hut, S. C., Hedlund, L., Leminen, M., & Leminen, A. (2019). Acquisition of L2 morphology by adult language learners. *Cortex*, *116*, 74-90. doi:10.1016/j.cortex.2019.01.012

Kochančikaitė, R., Shtyrov, Y., & Roll, M. (2020). *Lexical and decompositional processing in parallel? Evidence from Swedish, a pitch-accent language.* Manuscript in preparation.

Kutas, M., & Federmeier, K. D. (2014). Thirty years and counting: Finding meaning in the N400 component of the event related brain potential (ERP). *Annual Review of Psychology, 62*, 621-647. doi: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. doi:10.1126/science.7350657

Lau, E. F., Phillips, C., & Poeppel, D. (2008). A cortical network for semantics: (de)constructing the N400. *Nature Reviews Neuroscience*, *9*, 920-933. doi:10.1038/nrn2532

Liu, C. (2013). Just noticeable difference of tone pitch contour change for English- and Chinese-native listeners. *The Journal of the Acoustical Society of America*, *134*(4), 3011-3020. doi:10.1121/1.4820887

MacGregor, L. J., Difrancesco, S., Pulvermüller, F., Shtyrov, Y., & Mohr, B. (2015). Ultra-rapid access to words in chronic aphasia: The effects of intensive language action therapy (ILAT). *Brain Topography, 28*, 279-291. doi:10.1007/s10548-014-0398-y

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. doi:10.1038/ncomms1715

Maddieson, I. (2013). Tone. In M. S. Dryer, & M. Haspelmath (Eds.), *The world atlas of language structures online.* Leipzig: Max Planck Institute for Evolutionary Anthropology. Retrieved September 17, 2018, from http://wals.info/chapter/13

Malins, J. G., & Joanisse, M. F. (2012). Setting the tone: An ERP investigation of the influences of phonological similarity on spoken word recognition in Mandarin Chinese. *Neuropsychologia*, *50*, 2032-2043. doi:10.1016/j.neuropsychologia.2012.05.002

Maris, E., & Oostenveld, R. (2007). Nonparametric statistical testing of EEG- and MEG-data. *Journal of Neuroscience Methods*, *164*, 177-190. doi:10.1016/j.jneumeth.2007.03.024

Molinaro, N., Barber, H. A., Caffarara, S., & Carreiras, M. (2014). On the left anterior negativity (LAN): The case of morphosyntactic agreement. *Cortex*, 1-4. doi:10.1016/j.cortex.2014.06.009

Partanen, E., Leminen, A., de Paoli, S., Bundgaard, A., Kingo, O. S., Krøjgaard, P., & Shtyrov, Y. (2017). Flexible, rapid and automatic neocortical word form acquisition mechanism in children as revealed by neuromagnetic brain response dynamics. *NeuroImage*, *155*, 450-459. doi:10.1016/j.neuroimage.2017.03.066

Pelzl, E., Lau, E. F., Guo, T., & DeKeyser, R. (2019). Advanced second language learners' perception of lexical tone contrasts. *Studies in Second Language Acquisition, 41*, 59-86. doi:10.1017/S0272263117000444

Pelzl, E., Lau, E. F., Guo, T., & DeKeyser, R. (2020). Even in the best-case scenario L2 learners have persistent difficulty perceiving and utilizing tones in Mandarin: Findings from behavioural and event-related potentials experiments. *Studies in Second Language Acquisition*, 1-29. doi:10.1017/S027226312000039X

Qin, Z., & Mok, P. P.-K. (2011). Discrimination of Cantonese tones by Mandarin, English and French speakers. *Conference: the Psycholinguistic representation of tone conference (PLRT).*

Roll, M. (2015). A neurolinguistic study of South Swedish word accents: Electrical brain potentials in nouns and verbs. *Nordic Journal of Linguistics*, *38*(2), 149-162. doi:10.1017/S0332586515000189

Schremm, A., Novén, M., Horne, M., Söderström, P., van Westen, D., & Roll, M. (2018). Cortical thickness of planum temporale and pars opercularis in native language tone processing. *Brain and Language*, *176*, 42-47. doi:10.1016/j.bandl.2017.12.001

Shen, G., & Froud, K. (2019). Electrophysiological correlates of categorical perception of lexical tones by English learners of Mandarin Chinese: an ERP study. *Bilingualism: Language and Cognition*, *22*(2), 253-265. doi:10.1017/S136672891800038X

Shtyrov, Y., & Lenzen, M. (2017). First-pass neocortical processing of spoken language takes only 30 msec: Electropysiological evidence. *Cognitive Neuroscience*, *8*(1), 24-38. doi:10.1080/17588928.2016.1156663

So, C. K., & Best, C. T. (2014). Phonetic influences on English and French listeners' assimilation of Mandarin tones to native prosodic categories. *Studies in Second Language Acquisition*, *36*, 195-221. doi:10.1017/S0272263114000047

Söderström, P., Horne, M., & Roll, M. (2016). Stem tones pre-activate suffixes in the brain. *Journal of Psycholinguistic Research*, 1-10. doi:10.1007/s10936-016-9434-2

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. doi:10.3758/BF03192720

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. (2002). *Tone.* Cambridge: Cambridge University Press. doi:10.1017/CBO9781139164559

Yu, K., , L., Chen, Y., Zhou, Y., Wang, R., Zhang, Y., & Li, P. (2019). Effects of native language experience on Mandarin lexical tone processing in proficient second language learners. *Psychophysiology, 56:e13448.* doi:10.1111/psyp.13448

Yue, J., Bastiaanse, R., & Alter, K. (2014). Cortical plasticiy induced by rapid Hebbian learning of novel tonal word-forms: Evidence from mismatch negativity. *Brain & Language, 139*, 10-22. doi:10.1016/j.bandl.2014.09.007

Zhao, J., Guo, J., Zhou, F., & Shu, H. (2011). Time course of Chinese monosyllabic spoken word recognition: Evidence from ERP analyses. *Neuropsychologia, 49*, 1761-1770. doi:10.1016/j.neuropsychologia.2011.02.054

Supplementary material

Statistically significant results from the repeated measures Analysis of Variance for the positive posterior cluster at 400-600 ms.

Multiple comparisons are corrected with False Discovery Rate (FDR) corrections and effects only found in the posterior cluster are marked with an exclamation mark.

Abbreviations

LG	= Learner Group
TTG	= Target Tone Group
D	= Day
B	= Block
TL1s	= tonal L1 learner group
NTL1s	= non-tonal L1 learner group
H/F	= high/fall group
L/R	= low/rise group

MAIN EFFECTS, INTERACTIONS & PAIRWISE COMPARISONS	F	P	MEANS IN μV			
Tone Type	183.35	< .001	Contour	1.86	Level	2.21
Tone Type * Learner Group	4.43	.041				
TL1s: Tone Type	190.90	<.001	Contour	1.80	Level	2.20
NTL1s: Tone Type	48.12	<.001	Contour	1.93	Level	2.22
Tone Type * Day * Block * LG * TTG	3.89	.004				
H/F: Tone Type * Day * Block * LG	3.76	.006				
L/R: Tone Type * Day * Block * LG	1.34	.260				
H/F, D1: Tone Type * Block * LG	3.34	.015				
H/F, D1, B1: Tone Type * LG	0.34	.856				
H/F, D1, B2: Tone Type * LG	4.70	.041				
H/F, D1, B2, NTL1: Tone Type	26.22	<.001	Contour	2.01	Level	2.59
H/F, D1, B2, TL1: Tone Type	7.60	.019	Contour	2.25	Level	2.70
H/F, D1, B3: Tone Type * LG	3.09	.093				
H/F, D1, B4: Tone Type * LG	1.12	.301				
H/F, D1, B5: Tone Type * LG	2.13	.159				
H/F, D1, B6: Tone Type * LG	7.86	.010				
H/F, D1, B2, NTL1: Tone Type	8.35	.380	Contour	2.67	Level	2.79
H/F, D1, B2, TL1: Tone Type	34.52	<.001	Contour	2.13	Level	2.73
H/F, D2: Tone Type * Block* LG	2.60	.043				
H/F, D2, B1: Tone Type * LG	2.32	.142				
H/F, D2, B2: Tone Type * LG	0.33	.573				
H/F, D2, B3: Tone Type * LG	1.79	.195				
H/F, D2, B4: Tone Type * LG	4.63	.043				
H/F, D2, B2, NTL1: Tone Type	0.32	.585	Contour	2.09	Level	2.16
H/F, D2, B2, TL1: Tone Type	10.44	.008	Contour	1.81	Level	2.27
H/F, D2, B5: Tone Type * LG	1.33	.261				
H/F, D2, B6: Tone Type * LG	0.90	.354				

	Effect	Comparison	F	p	Level	M	Level	M
	Learning		30.76	<.001	Target	2.14	Control	1.93
	Tone Type * Learning * Day * TTG		6.37	.015				
	H/F: Tone Type * Learning * Day		1.05	.317				
	L/R: Tone Type * Learning * Day		8.89	.007				
	L/R, Contour: Learning * Day		0.01	.973				
	L/R, Level: Learning * Day		10.55	.004				
	L/R, Level, D1:	Learning	5.77	.025	Target	2.21	Control	2.03
	L/R, Level, D2:	Learning	17.79	<.001	Target	2.10	Control	1.75
!	Tone Type * Learning * Day * Learner Group		4.84	.033				
	Contour: Learning * Day * Learner Group		0.14	.708				
	Level: Learning * Day * Learner Group		5.05	.030				
	Level, D1: Learning * Learner Group		5.38	.025				
	Level, D1, TL1s:	Learning	19.36	<.001	Target	2.52	Control	2.21
	Level, D1, NTL1s:	Learning	1.46	.239	Target	2.46	Control	2.38
	Level, D2: Learning * Learner Group		1.79	.188				
	Level, D1, TL1s:	Learning	12.87	.002	Target	2.06	Control	1.86
	Level, D2, NTL1s:	Learning	1.71	.205	Target	2.17	Control	2.08
	Block		3.16	.013				
	multiple comparisons:	B1 vs. B2		.025	Block 1	2.14	Block 2	2.04
		B1 vs. B3		.048	Block 1	2.14	Block 3	2.02
		B1 vs. B5		.027	Block 1	2.14	Block 5	1.96
		B1 vs. B6		.042	Block 1	2.14	Block 6	1.99
	Day		24.10	<.001	Day 1	2.22	Day 2	1.85
	Day * TTG		5.55	.023				
	H/F: Day		44.76	<.001	Day 1	2.47	Day 2	1.93
	L/R: Day		2.31	.143	Day 1	1.96	Day 2	1.77
!	Day * Block * TTG		3.02	.031				
	Day 1: Block * TTG		2.15	.111				
	Day 2: Block * TTG		2.90	.038				
	Day 2, H/F: Block		3.01	.014				
	multiple comparisons:	B1 vs. B2		.011	Block 1	2.14	Block 2	1.93
		B1 vs. B3		.012	Block 1	2.14	Block 3	1.89
		B1 vs. B5		.021	Block 1	2.14	Block 5	1.78
		B1 vs. B6		.015	Block 1	2.14	Block 6	1.77
		B3 vs. B4		.017	Block 3	1.89	Block 4	2.08
	Day 2, L/R: Block		4.48	.004				
	multiple comparisons:	B1 vs. B3		.007	Block 1	1.96	Block 2	1.83
		B1 vs. B4		.013	Block 1	1.96	Block 3	1.71
		B1 vs. B5		.014	Block 1	1.96	Block 4	1.65
		B1 vs. B6		.017	Block 1	1.96	Block 5	1.75
	Day * Block * Learner Group * TTG		3.14	.027				
	TL1s: Day * Block * TTG		4.97	.005				
	TL1s, H/F: Day * Block		3.54	.038				
	TL1s, H/F, D1: Block		0.50	.662				
	TL1s, H/F, D2: Block		5.24	.020				
	multiple comparisons:	B1 vs. B2		.050	Block 1	2.12	Block 2	1.91
		B1 vs. B3		<.001	Block 1	2.12	Block 3	1.53
		B1 vs. B4		.047	Block 1	2.12	Block 4	1.81
		B1 vs. B5		<.001	Block 1	2.12	Block 5	1.64
		B1 vs. B6		<.001	Block 1	2.12	Block 6	1.62
		B2 vs. B3		<.001	Block 2	1.91	Block 3	1.53
		B2 vs. B5		.044	Block 2	1.91	Block 5	1.64
		B2 vs. B6		<.001	Block 2	1.91	Block 6	1.62
		B3 vs. B4		.043	Block 3	1.53	Block 4	1.81
	TL1s, L/R: Day * Block		2.57	.081				
	NTL1s: Day * Block * TTG		0.99	.402				