

Word tones cueing morphosyntactic structure: Neuroanatomical substrates and activation time-course assessed by EEG and fMRI



Mikael Roll ^{a,*}, Pelle Söderström ^a, Peter Mannfolk ^b, Yury Shtyrov ^c, Mikael Johansson ^d, Danielle van Westen ^e, Merle Horne ^a

^a Department of Linguistics and Phonetics, Lund University, Sweden

^b Department of Medical Radiation Physics, Clinical Sciences, Lund University, Sweden

^c Center of Functionally Integrative Neuroscience, Institute for Clinical Medicine, Aarhus University, Denmark

^d Department of Psychology, Lund University, Sweden

^e Department of Diagnostic Radiology, Clinical Sciences, Lund University, Sweden

ARTICLE INFO

Article history:

Received 13 January 2015

Revised 24 June 2015

Accepted 18 July 2015

Keywords:

Word accent

Lexical tone

Morphology

Grammar

ERP

fMRI

Superior temporal gyrus

Inferior frontal gyrus

ABSTRACT

Previous studies distinguish between right hemisphere-dominant processing of prosodic/tonal information and left-hemispheric modulation of grammatical information as well as lexical tones. Swedish word accents offer a prime testing ground to better understand this division. Although similar to lexical tones, word accents are determined by words' morphosyntactic structure, which enables listeners to use the tone at the beginning of a word to predict its grammatical ending. We recorded electrophysiological and hemodynamic brain responses to words where stem tones matched or mismatched inflectional suffixes. Tones produced brain potential effects after 136 ms, correlating with subject variability in average BOLD in left primary auditory cortex, superior temporal gyrus, and inferior frontal gyrus. Invalidly cued suffixes activated the left inferior parietal lobe, arguably reflecting increased processing cost of their meaning. Thus, interaction of word accent tones with grammatical morphology yielded a rapid neural response correlating in subject variability with activations in predominantly left-hemispheric brain areas.

© 2015 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

1. Introduction

Understanding everyday speech requires identification of speech sounds, word stems and affixes, as well as integrating this information in grammatical structures – all at rates of up to 6–7 syllables/second (Levelt, 1989). This remarkable performance would not be possible without the help of cues in speech that continuously provide information about the upcoming structure of words and clauses (Cutler, Dahan, & Donselaar, 1997; Roll, Horne, & Lindgren, 2011). Many languages use intonation to cue grammatical structure at the clause level. This type of interaction is known to give rise to inter-hemispheric signal exchange due to the left hemisphere's specialization in modulating grammar and the right hemisphere's dominance in regulating intonation (Friederici & Alter, 2004; Sammler, Kotz, Eckstein, Ott, & Friederici, 2010). However, while grammar processing is believed

to rely on left-lateralized perisylvian regions (Marslen-Wilson & Tyler, 2007), laterality of tonal processing is less clear-cut; not all kinds of tonal information are considered to have a mainly right-hemispheric substrate. Crucially, word tones in Chinese or Thai, which serve categorical lexico-semantic distinctions, have been found to increase activation in areas of the left superior temporal gyrus (Xu et al., 2006). If both word tones and grammar show a left hemisphere bias, word tones cueing grammatical affixes should give rise to tone-grammar association activations concentrated in the left hemisphere. This can be tested in Swedish, where listeners use word tones on stems to unconsciously predict which suffix words will have (Roll, Horne, & Lindgren, 2010). Brain areas responsible for morphological prediction have to be rapidly activated, since in disyllabic words, prediction occurs from one syllable to the next, which limits the time frame for this process to ~150 ms at fast speech rates (Roll, Söderström, & Horne, 2013). To address the spatio-temporal dynamics of these fundamental processes connecting prosody to grammar, the present study investigated the neural substrates that enable rapid associations between stem tone and suffix using a combination of temporally

* Corresponding author at: Center for Languages and Literature, Lund University, Box 201, 22100 Lund, Sweden.

E-mail address: mikael.roll@ling.lu.se (M. Roll).

and spatially resolved neuroimaging methods. Electroencephalographic (EEG) recordings established a timeline of brain responses to stem tones and suffixes, and functional magnetic resonance imaging (fMRI) was used to identify the location of the brain areas involved in the processing.

1.1. Morphosyntactic word tones

Swedish (and related Norwegian) have long been known to have word tones similar to those in e.g. Chinese, called “word accents” (Bruce, 1977; Chao, 1976). However, in Swedish and Norwegian, the tone that is realized on a word’s stem depends on which suffix is attached to the stem (Riad, 2012; Rischel, 1963). Therefore, the Swedish stem *hatt* in *hatt + en* (‘the hat’) has a low tone, known as “Accent 1” (Fig. 1, A1), whereas the same stem in *hatt + ar* (‘hats’) has a high tone, or “Accent 2” (Fig. 1, A2), due to the suffix difference.

Word accent tones have a clear function in facilitating rapid word-processing in Swedish (Roll et al., 2010). Thus, native speakers use the association between stem tone and suffix to predict which ending a word will have already when hearing the stem. This has been seen in increased response times for judging the meaning of suffixes that were invalidly cued by the wrong stem tone (Söderström, Roll, & Horne, 2012). Electrophysiological studies have shown a difference in the processing of Accent 1 and Accent 2 starting around 140 ms from tone onset (Roll et al., 2010, 2013). This differential effect has previously been thought to reflect an increase in neural activity for Accent 2 (expressed as a positivity on the scalp surface) due to the high tone’s relative auditory salience compared to the low Accent 1 tone (Roll et al., 2010). However, the electrophysiological effect has only been seen to appear if the tonal contrast is realized within existing words (Roll et al., 2013), suggesting that it might reflect a process involved in the predictive function of word accents (signaling upcoming suffixes) rather than indexing tonal salience. In human vocalizations that do not include meaningful lexical or syntactic information (humming), the auditory salience of the high Accent 2 tone has instead produced an increase in the auditory N1 component (Roll et al., 2013).

Accent 1 is more useful than Accent 2 for predicting its related suffixes, since it is associated with a well-defined set of endings, whereas Accent 2, in addition to connecting to a set of suffixes, occurs in compound words as well, e.g. *hattband* ‘hat band.’ Accordingly, validly cued Accent 1 suffixes have yielded shorter response times than Accent 2 suffixes (Roll et al., 2013;

Söderström et al., 2012). Therefore, it seems likely that the differential effect previously found for word accents should be interpreted as a negativity for Accent 1 indexing greater preactivation of memory traces of the suffixes associated with the tone.

1.2. Neural substrates for processing tone and grammar

Whereas the above studies have given some limited information on the EEG time-course for tone-suffix interaction, the neural substrates underlying this interaction remain unknown. However, based on studies in Thai, Chinese, and non-tonal languages like English or German, morphosyntactic tone could be expected to engage a number of brain areas. One strong candidate is the left inferior frontal gyrus (LIFG), which is assumed to subservise grammar processing (Marslen-Wilson & Tyler, 2007). Processing of word structure (morphology) has been seen to specifically involve the ventral part of LIFG, Brodmann area (BA) 47 (Koester & Schiller, 2011; Tyler, Marslen-Wilson, & Stamatakis, 2005). Dichotic listening studies indicate that native speakers of tone languages such as Mandarin and Norwegian engage the left hemisphere more than nonnative speakers when discriminating word tones (Moen, 1993; Wang, Jongman, & Sereno, 2001; Wang, Sereno, Jongman, & Hirsch, 2003). Studies involving brain-damaged patients have confirmed the dominance of the left hemisphere in word tone processing in these languages (Hughes, Chan, & Su, 1983; Moen & Sundet, 1996; Naeser & Chan, 1980; Packard, 1986) as well as in Thai (Gandour et al., 1992), Toisanese Chinese (Eng, Obler, Harris, & Abramson, 1996), Cantonese (Yiu & Fok, 1995), and Shona, a Bantu language in which tone is conditioned by word structure in a manner similar to Swedish and Norwegian (Kadyamusuma, De Bleser, & Mayer, 2011).

Brain imaging studies have found activation in the left frontal lobe and inferior parietal lobe for tasks involving active discrimination between lexical word tones (Gandour et al., 2000, 2003, 2004; Klein, Zatorre, Milner, & Zhao, 2001). A problem when interpreting results of tone discrimination in languages with lexical tone is that there is a confound between tonal and lexical analysis. In order to isolate the prelexical processing of tones, Xu et al. (2006) let Thai speakers listen to Thai tones, Mandarin tones, and Thai tones superimposed on Chinese syllables. They found overlapping activation for tone processing in both known words and unknown words in the temporal plane of the left superior temporal gyrus (STG), involving BA 22, 41, and 42. This area is convergent with that associated with the analysis of segmental speech sounds (consonants and vowels) (Graves, Grabowski, Mehta, & Gupta, 2008). In other words, it would seem that familiar word tones are processed like any other phonologically distinctive sound in the STG even without any associated meaning. The activity in frontal and parietal cortex in previous studies might have been due to selection related to the tone-discrimination task and processing the word meaning. Still, no neuroimaging evidence is available for the predictive morphosyntactic processes of the kind that characterize tone-suffix associations such as those in Swedish.

1.3. Present study

We used fMRI and EEG to comprehensively investigate morphosyntactic tone processing in the brain. EEG has excellent temporal resolution, and fMRI is superior in detecting sources of brain activity. Although electrophysiological and metabolic measures are sensitive to different temporal and spatial scales, using both methods in the same paradigm on the same participants allowed us to conjecture possible relationships between temporally and spatially resolved effects. We hypothesized that the brain treats the difference between word accent tones in Swedish as a phonological distinction. Therefore, we expected the increased

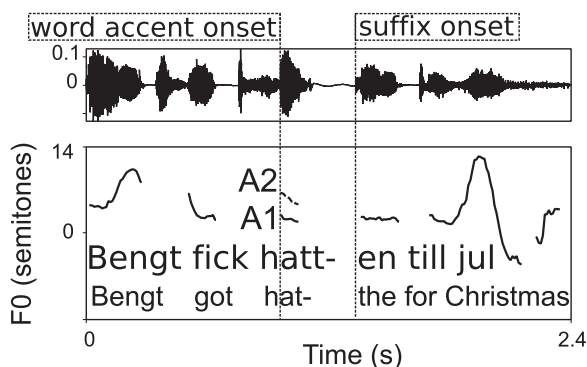


Fig. 1. Example of a stimulus sentence. Acoustic waveform and fundamental frequency (F0) are shown. The solid F0 line at *hatt* ‘hat’ represents the low Accent 1 tone associated with the definite singular ending *-en*. The broken line indicates what the corresponding high Accent 2 tone would have been had the ending been plural *-ar*.

neural activity found for Accent 1 as compared to Accent 2 in Roll et al. (2010) to stem from brain areas involving BA 22 and BA 41/42 in the STG, similar to what has been previously found for phonological distinctions. Due to its greater predictive value, Accent 1 would also be expected to show increased preactivation of its associated suffixes. That could involve activity in BA 47 in the left inferior frontal gyrus, as has previously been found for morphological processing (Koester & Schiller, 2011; Tyler et al., 2005). Suffix activation should ideally occur within the first 150 ms in order for word accents to be useful cues at faster speech rates. We hypothesized that the previously found ERP (event-related potential) effect beginning at ~140 ms after tone exposure indexes tone analysis and suffix activation. To see whether Accent 1 or Accent 2 increased the overall activity more and at which latencies, we performed a global RMS analysis.¹ Correlation between participant mean BOLD and ERP effects could give a hint as to the timeline of anatomically resolved effects. Although both effects are related to neural activity, care should nevertheless be taken in the interpretation of the relation between EEG and BOLD signals, since they may sometimes stem from different neural sources (Ritter & Villringer, 2006).

Invalidly cued suffixes have previously been found to produce increased positivity between 400 and 600 ms after suffix onset (Roll et al., 2010, 2013). This has been interpreted as a P600-like effect, showing reprocessing of the incorrect word form. Word-form processing would be thought to correlate with increased BOLD activity in temporal areas. However, if participants predict an upcoming suffix based on the word accent cue, the P600 could also reflect the correction of a failed prediction about number (e.g. 'singular' to 'plural'). This being the case, invalidly cued suffixes might increase activity in the inferior parietal lobe (IPL), known to be involved in number processing (Hirsch, Moreno, & Kim, 2001; Piras & Paola, 2009).

To be able to relate brain-imaging data to behavioral findings, participants performed a number-decision task. They were instructed to decide as quickly as possible whether the person in the sentence got "one" or "several" things by left- or right-hand button-press. Hand was counterbalanced within participants. To avoid the possibility that effects depended on participants focusing on the suffix due to the task, a control task was used in half of the blocks. This task was to alternately press right- and left-hand buttons as quickly as possible when the whole sentence ended, and thus did not require any attention to the suffix. Response times were measured for the number-decision task, whereas EEG and fMRI data were collapsed over both tasks. Hence, the experiment had two experimental factors, word accent (1, 2), and validity (valid, invalid), giving four conditions. Validity was relevant only for the suffix point measurements. For the word accent point measurements, therefore, ERPs were collapsed over validity. Due to its poor time resolution, the BOLD signal to word accents could have been affected by validity. Therefore, rather than collapsing the validity condition for the fMRI, a conjunction analysis was performed, where only significant activations coinciding in both valid and invalid conditions of the word accent contrast were taken into account.

2. Materials and methods

2.1. Participants

Eighteen right-handed native speakers of Central Swedish, mean age 25.3 years, $SD = 5.3$, 8 women, participated in the study.

¹ While increased gRMS can reflect increased neural activity, it can also stem from greater neural synchronization or release from inhibitory influence.

All were undergraduate students at Lund University. The local ethics board approved the study.

2.2. Procedure and stimuli

Sixty different sentences per condition, 240 in total, were presented in 4 blocks in pseudorandomized order with SOA (stimulus onset asynchrony) jittered between 4 and 8 s using Optseq2 (Dale, 1999). Order of stimuli, tasks, and hand-response associations was counterbalanced across participants. fMRI sessions took place a few days after the ERP sessions, using the same paradigm version within participants to make the settings as similar as possible. All sentences had the same syllable structure:

Kurt fick hatten/hattar till jul
'Kurt got the-hat/hats for Christmas'

Carrier sentences with prosodic focus on the last prepositional phrase (*till jul* 'for Christmas' in the example) were used in order to avoid focus on the critical object noun, i.e. *hatten/hattar* 'the hat/'hats' in the example, since focus interacts with word accents (Bruce, 1977), and thus would make results more difficult to interpret. Stimulus nouns containing two syllables with voiceless stops at the boundary between stem and suffix (either definite singular or plural, e.g. *hatt-en/ar* hat-the/s) were chosen for ease of splicing. The critical noun was also always separated by a voiceless stop from the surrounding sentence material (*fick* 'got' and *till* 'for' in the example). A male Central Swedish speaker recorded the sentences in an anechoic chamber. Critical words were extracted from the carrier sentence and were cut between stem and suffix in order to create combinations for cross-splicing stems with validly and invalidly cued suffixes. The intensity was normalized separately over Accent 1 stems, Accent 2 stems, and suffixes. The stem/suffix fragments were spliced in valid and invalid combinations, which were then spliced back into the carrier sentences. The first part of the carrier sentences (*Kurt fick...* 'Kurt got...') measured 925 ms in duration, $SD = 92$ ms. Stems (*hatt-* 'hat-') were on average 428 ms long for Accent 1, $SD = 60$ ms, and 424 ms long for Accent 2, $SD = 67$. Accent 1 suffixes (-*en* 'the,' singular definite) were 241 ms, $SD = 24$ ms, and Accent 2 suffixes (-*ar/-er* '-s,' plural) were also 241 ms, $SD = 24$ ms. The duration of the part of the carrier sentence following the suffix (*till jul* 'for Christmas') was 873 ms, $SD = 68$ ms. The parts preceding and following the critical noun were identical across conditions, and were in half of the cases taken from the singular recording, and in half, from the plural recording. The low Accent 1 tone was 2.71 semitones (st) at vowel onset, $SD = 0.86$, and fell to 1.30 st, $SD = 0.52$, during 137 ms, $SD = 15$. The corresponding high Accent 2 tone was 6.76 st, $SD = 1.53$ st, falling to 3.08 st, $SD = 2.25$, with a duration of 117 ms from vowel onset to offset, $SD = 27$. The cross-splicing created a balanced design where the same stems and suffixes appeared in valid and invalid combinations, thereby ruling out purely acoustic explanations for any effects.

2.3. Time-locking points

The time-locking point for the word accent contrast in ERP and fMRI was stem-vowel onset in the first syllable of critical words, e.g. *a* in *hatten* 'the hat' ("word accent onset" in Fig. 1). This is the point where fundamental frequency starts to differ between conditions. For the suffix contrast, ERP, fMRI, and reaction times were time-locked to onset of the syllable containing the suffix i.e. the burst of [t] in *hatten/hattar* 'the hat/hats' ("suffix onset" in Fig. 1). It is possible to distinguish plural from singular suffixes at this point due to co-articulation between the consonant (here [t]) and the suffix vowel ([e] or [a] in *-er* and *-ar*).

2.4. Electroencephalography

Participants sat in front of a computer screen listening to stimuli via loudspeakers. A 32-channel EasyCap and a Synamps 2 amplifier recorded the EEG at a sampling rate of 250 Hz. A band-pass filter with cutoff frequencies 0.05–70 Hz was used online, and a 30 Hz low-pass filter was applied offline. Impedances were kept below 5 kΩ. A centrofrontal electrode (FCz) was used as online reference, and data were re-referenced offline to an average reference.

Epochs starting at 200 ms before the word accent and suffix points, and ending 1000 ms thereafter were extracted. Thirty epochs were extracted per subject and condition. A 200 ms pre-stimulus time window was used for baseline correction. Figs. 2 and 3 show time-windows where relevant effects are visible. Epochs with voltage exceeding ±100 μV after compensation for eye artifacts using independent component analysis (Jung et al., 2000) were discarded, leaving 28.1 epochs per condition (SD = 2.6).

Reference-free, global root mean squares (gRMS) (Lehmann & Skrandies, 1980) were calculated for the individual ERPs. To obtain some expectation of where peaks could be found, we analyzed gRMS for the data from a previous study investigating a similar contrast (Roll et al., 2010), and observed peaks at 140, 180 and 260 ms (Supplementary Material). Where visual inspection and previous research suggested an effect in the ERPs, we inspected 30 ms time-windows surrounding gRMS peaks. Average gRMS of all unrejected epochs were submitted to repeated measures ANOVAs. With the goal of replicating the previously found negativity for Accent 1, an ANOVA of average ERPs was also performed with additional topographical factors anterior-posterior (antpost) and hemisphere (hem), corresponding to regions left anterior (F7, F3), right anterior (F4, F8), left central (T7, C3), right central (C4,

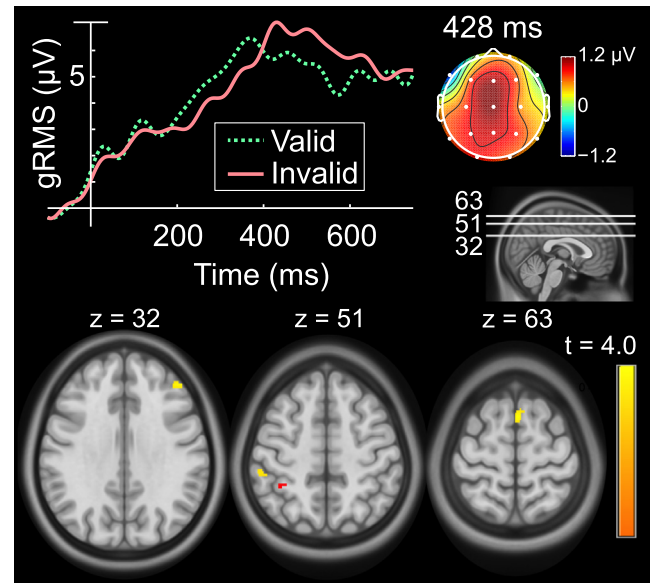


Fig. 3. Neural activity produced by suffixes. Top: Suffixes that were preceded by the wrong word accent (invalidly cued) yielded increased global root mean square (gRMS) values peaking at 428 ms (left). This corresponded to a positivity in the ERPs, seen in a topographical map (mid). The effect was stronger in Accent 1 words, as seen in a word accent × suffix interaction (right). Bottom: T maps from fMRI showed increased activity for invalidly cued suffixes in several areas, $p < 0.001$ (uncorrected). The Invalid > Valid contrast in an area in the inferior parietal lobe (red color) correlated with the gRMS effect in terms of subject variability.

T8), left posterior (P7, P3), and right posterior (P4, P8). The time window for ERP analysis of word accent effects was 136–280 ms after F0 onset. This was somewhat earlier than the 200–300 ms

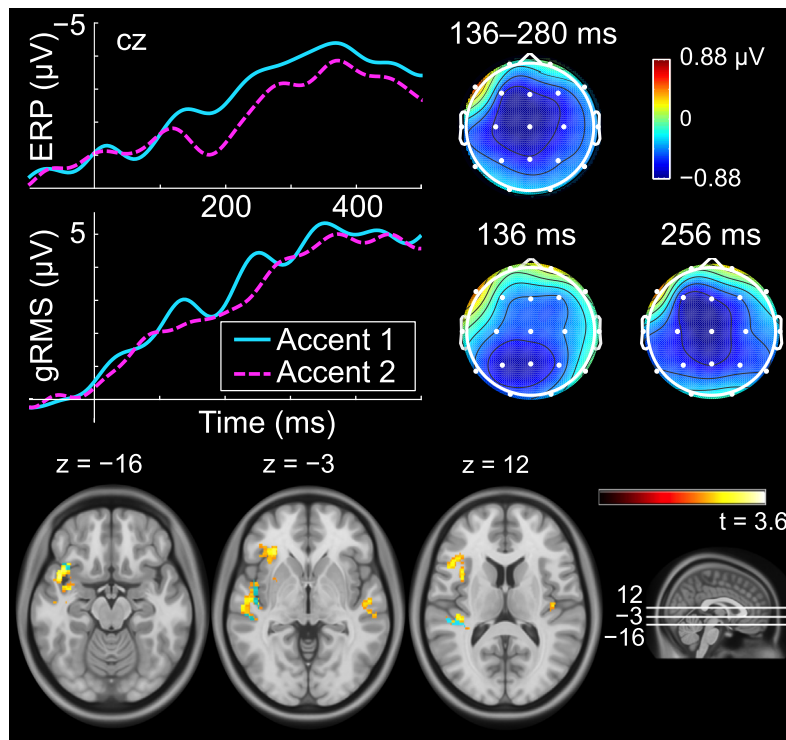


Fig. 2. Neural activation of word accents. Top: ERP waveform at central electrode showing increased negativity for Accent 1 between 136 and 280 ms. Subtraction map indicates left-lateralized topographical distribution. Mid: Global root mean squares (gRMS) confirm increased neural activity for Accent 1 at 136 and 256 ms, with left posterior and left anterior topographical distribution. Bottom: Red-yellow-white color transition represents t values for Accent 1 > Accent 2 contrast in fMRI, $p < 0.05$, FDR-corrected. Green-blue color specifies areas in primary auditory cortex, superior temporal gyrus, and inferior frontal gyrus where there was correlation in subject variability of the Accent 1 > Accent 2 contrast and the same contrast in gRMS.

time window used in Roll et al. (2010). In this previous study, onset latency was not closely investigated. However, closer scrutiny of the previous data showed a stable ERP difference already at 136 ms. This indicates an onset latency of ~ 140 ms for the word accent effect in the previous study. If gRMS is assumed to reflect the onset of important changes in neural activity, the first peak found in the earlier data suggested that a time window onset at around 140 ms would be pertinent. At suffix onset, the factor validity (valid, invalid) was added. A gRMS peak before the observed 450–700 ms time window for P600 was expected (Roll et al., 2010). Greenhouse–Geisser correction was used when applicable. All and only significant effects are reported.

2.5. fMRI

A Siemens Magnetom Skyra 3.0T was used for the acquisition of MRI data from the same experimental participants using a 32-channel head coil. A gradient-echo EPI pulse sequence produced T2* contrast images for BOLD data (TR = 2000 ms, TE = 30 ms, flip angle = 90°, field of view = 192 mm, matrix size = 64 × 64, 33 slices, slice thickness = 3 mm). A T1-weighted MPRAGE pulse sequence was used for overlay of statistical results. Furthermore, a T2-weighted FLAIR pulse sequence was used in order to exclude pathology. Preprocessing and statistical analysis were performed with SPM8 software (Wellcome Department of Cognitive Neurology, <http://www.fil.ion.ucl.ac.uk/spm>). Pre-processing included motion correction, slice timing correction, normalization to standard MNI space and smoothing (6 mm isotropic Gaussian kernel) to fulfill the assumptions of Gaussian random field theory (Worsley, Poline, Vandal, & Friston, 1995). For normalization, the SPM8 EPI template was used, which is based on a standard Montreal Neurological Institute (MNI) space (Ashburner & Friston, 1999; Friston et al., 1995).

Beta values from each condition and subject were calculated timed to voice onset in critical words for tone effects and to suffix onset for suffix effects. Beta values were estimated in a fixed-effects first-level analysis using event-related design. High-pass filter was 128 s long. In the second-level analysis, beta values entered a full factorial ANOVA with 2 dependent levels: word accent and validity. To avoid increased influence of specific invalid tone-suffix combinations, we performed a conjunction analysis between words containing validly and invalidly cued suffixes for tone effects. The extent threshold was 10 voxels. FDR-corrected t maps are shown in Fig. 2. A threshold of $p < 0.025$ rather than $p < 0.05$ was used to compensate for increased power of assuming a global null hypothesis. For suffix effects, valid-invalid contrasts were calculated collapsing Accent 1 and Accent 2 as well as singular and plural suffix forms. This was done to avoid influence of individual word accents or suffixes, and thus isolate brain areas specifically sensitive to invalidly cued suffixes. The threshold was lowered to uncorrected $p < 0.001$ in Fig. 3 due to lack of effects using FDR correction. Subregions of the activation of the word accent contrast were also extracted, and Brodmann areas (BA) were defined by the Talairach Daemon database (Lancaster, Summerlin, Rainey, Freitas, & Fox, 1997) in the PickAtlas Software toolbox (Maldjian, Laurienti, Kraft, & Burdette, 2003).

2.6. EEG-fMRI correlations

To see whether BOLD effects and gRMS effects were related, so that an individual increase in a BOLD effect corresponded to an individual increase in a gRMS effect, we calculated the correlation coefficient for participant means (Bland & Altman, 1996). For each participant, MarsBaR extracted the average beta value of the Accent 1–Accent 2 subtraction in each region where a significant “Accent 1 effect” had been found. Whole regions of focal BOLD

effects and subregions of more extensive BOLD effects were used. The Accent 1 effect was correlated with the corresponding average subtraction of Accent 1–Accent 2 at significant peaks in the gRMS, both peak upstroke and downstroke (3–6 samples before and after the peak). The same procedure was followed for “validity effects” (invalid–valid subtraction). One-tailed t tests tested significance of correlation. Bonferroni-corrected p values are reported. Both BOLD effects and gRMS were correlated with the response time contrast between validly and invalidly cued Accent 1-associated suffixes, showing an advantage for Accent 1 as a suffix cue.

3. Results

3.1. Behavior

Suffixes that had been validly cued by the correct word accent tone were processed faster than invalidly cued suffixes, $F(1, 17) = 16.73$, $p < 0.001$. Accent 1 was confirmed to be a stronger predictor for its associated suffixes than Accent 2, as indicated by a validity × word accent interaction, $F(1, 17) = 11.84$, $p < 0.003$. In Accent 1 words, valid suffixes were processed significantly faster than invalid suffixes (633 vs. 694 ms, $F(1, 17) = 21.33$, $p < 0.001$). The temporal advantage for validly cued suffixes in Accent 2 words (659 vs. 672 ms) was not significant, $F(1, 17) = 1.81$, $p = 0.196$.

3.2. Stem tone effects in the brain

In the brain potentials, Accent 1 yielded increased negativity as compared to Accent 2 between 136 and 280 ms with a left-lateralized distribution. This effect was seen in a word accent × antpost × hem interaction, $F(1, 17) = 3.87$, $p = 0.039$, a word accent × antpost interaction in the left hemisphere, $F(1, 17) = 3.91$, $p = 0.048$, and effects of word accent at left central, $F(1, 17) = 19.03$, $p < 0.001$, and left posterior sites, $F(1, 17) = 8.74$, $p = 0.009$. To obtain further understanding of the differential effect, global RMS were calculated. Peaks were found for Accent 1 at 136 ms, $F(1, 17) = 4.76$, $p = 0.043$, corresponding to a left posterior ERP distribution, and at 256 ms, $F(1, 17) = 9.23$, $p = 0.007$, with a left anterior distribution (Fig. 3). We hypothesized that the effect for Accent 1 was due to its greater use as a cue for the suffix. Speaking in favor of this hypothesis is a correlation in subject variability of the average 136 ms-peak in the gRMS Accent 1 > Accent 2 contrast and the behaviorally determined response time advantage (61 ms on average) for valid over invalid Accent 1-associated suffixes ($r = 0.513$, $p = 0.030$, peak downstroke).

In the fMRI, Accent 1 produced left-hemispheric activations in primary auditory cortex (A1, BA41), superior temporal gyrus (STG, BA22), mid temporal gyrus (BA21), the temporal pole (BA38), and the inferior frontal gyrus (IFG, BA47, and BA45) and frontal operculum (BA44). To assess the left-lateralization of these regions, we created Rols for their right hemisphere homologues. We then submitted the BOLD beta values to hemisphere × word accent repeated measures ANOVAs. Regions in IFG, frontal operculum, and the temporal pole showed a significant hemisphere × word accent interaction (Table 1). Follow-up ANOVAs to a marginal interaction in A1 (Table 1) revealed more than twice the effect size for word accent in the left hemisphere, $\eta_p^2 = 0.594$, $F(1, 17) = 24.85$, $p < 0.001$, as compared to the right hemisphere, $\eta_p^2 = 0.283$, $F(1, 17) = 6.72$, $p = 0.019$. Similarly, a marginal interaction in BA 21 (Table 1) showed a significant effect in the left, $F(1, 17) = 16.63$, $p = 0.001$, $\eta_p^2 = 0.495$, but not in the right hemisphere, $F(1, 17) = 3.06$, $p = 0.098$, $\eta_p^2 = 0.153$.

Subject variability in the average difference between Accent 1 and Accent 2 in global RMS following the early peak at 136 ms strongly correlated with the subject variability of the same average

Table 1

BOLD effects of the Accent 1 > Accent 2 contrast. Activations were found in bilateral superior temporal gyrus (STG), as well as left Heschl's gyrus (HG), middle temporal gyrus (MTG), temporal pole (TP), inferior frontal gyrus (IFG), and frontal operculum (FO). Word accent \times hemisphere interaction is shown for activated Brodmann areas.

Brodman area	Structure	Peak MNI coordinates	Size (voxels)	Peak <i>t</i>	WA \times hem <i>F</i> (1, 17)
41	STG, HG	−40, −30, 14	40	3.07**	3.29 [†]
22	STG	−46, −22, 2	63	3.12**	0.82
21	STG, MTG	−40, −8, −12	23	2.93**	4.09 [†]
38	TP	−44, 18, −22	58	3.53**	6.81*
47	IFG	−30, 26, −4	27	3.15**	9.37**
44	FO	−50, 18, 10	22	2.93**	8.70**
45	IFG	−46, 20, 12	22	2.75**	7.59*
13	INS	−38, 8, 14	40	3.37**	2.09
22 (right)	STG	46, −20, 4	27	2.79**	0.82

* $p < 0.05$.

** $p < 0.01$ (FDR-corrected in *t* tests).

[†] $p < 0.10$.

Table 2

Brain regions where subject variability in average BOLD activation correlated with subject variability in average ERP effects (gRMS) at 136 and 256 ms.

		BOLD activation		
		BA 41	BA 22	BA 47
ERP	136 ms	0.830***	0.646**	0.554*
	256 ms	–	–	0.498*

* $p < 0.05$.

** $p < 0.01$.

*** $p < 0.001$ (Bonferroni-corrected).

BOLD contrast in left A1, STG, and IFG regions (Table 2). In the same manner, the global RMS difference following the later peak at 256 ms correlated significantly in subject variability with the BOLD contrast in the left IFG (BA47). A correlation was also found for subject variability in average activation of A1 and average response time advantage for validly cued Accent 1 suffixes, $r = 0.547$, $p = 0.009$. No further significant correlations were found. Left insular cortex (BA13) and right STG (BA22) were also activated for Accent 1 > Accent 2. There were no areas with significant activation advantage for Accent 2 over Accent 1.

3.3. Suffix effects in the brain

The ERPs for validly and invalidly cued suffixes resulted in similar ERP waveforms, involving a late positive peak (Fig. 3). However, the gRMS peak had a longer latency (428 ms) for invalidly cued suffixes than for validly cued (368 ms). This timing difference is likely due to the fact that validly cued suffixes were processed faster. The peak for invalidly cued suffixes also had greater amplitude, $F(1, 17) = 11.99$, $p = 0.003$. As mentioned above, this effect has previously been interpreted as a 'P600,' indexing restructuring of the incorrect word forms. The invalid > valid contrast in the fMRI yielded an increased BOLD signal in the left inferior parietal lobe (IPL, BA40, Fig. 3) when the threshold was lowered from FDR-corrected to uncorrected $p < 0.001$ (Table 3). The activation was divided into a lateral spot, and a more central spot. Subject variability in the average P600 effect correlated with subject variability in the average BOLD signal in central part of IPL, $r = 0.551$, $p = 0.018$. Bilateral supplementary motor areas (BA 6) and the right middle frontal gyrus (BA 9) were also activated for invalidly cued suffixes, but did not correlate with the gRMS effects.

Table 3

BOLD effects of invalid suffixes > valid suffixes. Activations were found in left inferior parietal lobe (IPL), right middle frontal gyrus (MFG), and bilateral superior frontal gyrus (SFG).

Brodman area	Structure	Peak MNI coordinates	Cluster size (voxels)	Peak <i>t</i>
40	IPL	−54, −34, 48	22	3.48*
40	IPL	−36, −44, 48	14	3.37*
9 (right)	MFG	44, 38, 34	21	4.04**
6 (bilateral)	SFG	2, 12, 62	46	4.02**

* $p < 0.0005$.

** $p < 0.00005$ (uncorrected).

4. Discussion

Previous studies have found interhemispheric communication in the processing of tonal (right-hemisphere modulated) and grammatical (left-hemisphere modulated) information (Friederici & Alter, 2004; Sammler et al., 2010). The present study found rapid neural response and activation of brain areas concentrated to the left hemisphere for word tones interacting with inflectional suffixes. Based on the combined EEG-fMRI results, we propose the following time-course and cortical areas involved in the association between tone and grammatical suffixes. Tones are distinguished in the primary auditory cortex (BA41) at around 140 ms after tone onset, and immediately activate a phonological representation in the STG (BA22). The areas activated in the left temporal lobe are similar to those found for pre-attentive word tone processing in Thai listeners (Xu et al., 2006) and processing of speech sounds (Graves et al., 2008). Thus, Scandinavian word tones seem to have a neural representation more on a par with Chinese and Thai word tones than with sentence-level intonation, which has been observed to be more right-hemisphere biased (Gandour et al., 2003). This is supported by previous behavioral (Moen, 1993) and brain lesion findings (Moen & Sundet, 1996), and probably reflects the fact that just as Chinese tones are associated with specific words, Scandinavian tones are associated with specific affixes in the mental lexicon (Riad, 2012). A negativity with a topographical and temporal distribution similar to the ERP effect for Accent 1 has previously been found for non-attentive processing of Thai tones (Kaan, Barkley, Bao, & Wayland, 2008). This indicates that Scandinavian word accent tones and Thai tones might have comparable neural substrates and have a similar time course for phonological processing.

Although word accent tones are realized on word stems, it is the suffix that determines which tone the stem should carry. This makes it possible for listeners to predict the upcoming suffix upon hearing the tone. Immediately upon identification (at 136 ms after exposure), stem tones trigger a suffix prediction engaging the IFG (BA47). Activation of BA47 is in line with previous results for affix processing (Koester & Schiller, 2011; Tyler et al., 2005), and the latency of the ERP effect correlating in subject variability is similar to findings of activation of memory traces for grammatical affixes (Pulvermüller & Shtyrov, 2006; Shtyrov & Pulvermüller, 2002). It is likely that both tones use the same brain regions since Accent 2 was not found to engage any additional brain areas as compared to Accent 1. However, since Accent 1 is a stronger suffix predictor than Accent 2, it leads to increased activation.

The suffix prediction initiates morphological processing associated with the IFG peaking 120 ms later, at 256 ms. This is reminiscent of the N280, a negativity found for closed class (grammatical) words (Neville, Mills, & Lawson, 1992). In a manner similar to that associated with the effect of word accents, the N280 has been observed to increase for items with higher predictive value (Brunelliere, Hoen, & Dominey, 2005). A negativity within the same

time range has also been found for morpheme processing in Arabic (Boudelaa, Pulvermüller, Hauk, Shtyrov, & Marslen-Wilson, 2009).

Left lateralization for word accent processing was clearest in frontal BOLD effects, which are likely to be more related to suffix activation than tone processing. Furthermore, in A1 and the mid temporal lobe, left hemisphere effect sizes were more than twice as large as those in the right hemisphere, whereas the superior temporal gyrus (BA22) was bilaterally activated.

The morphological processing initiated by the tonal information might further activate a representation of the suffix's meaning (in the present study 'singular' or 'plural'). Invalidly cued suffixes challenge the number activation leading to increased activity in the inferior parietal lobe, an area which has been found to be involved in number processing (Hirsch et al., 2001; Piras and Paola, 2009). The parietal rather than temporal activation suggests that it is the activated meaning of the suffix that is reprocessed rather than its form.

Lastly, all correlations with subject variability in average BOLD were found for average gRMS peak downstrokes. The lack of upstroke correlation indicates that the gRMS peak shows the onset of the relevant neural events.

5. Conclusions

Left-hemispheric brain regions involving primary auditory cortex, STG, and IFG were found to be involved in the association of stem tones with grammatical suffixes. This left-hemispheric dominance for processing word accent-suffix connections differs from the previously found interhemispheric processing of tonal and grammatical information. Tones on stems increased brain potentials at 136 ms after exposure. The activation is likely to correspond to phonological analysis in primary and secondary auditory cortices as well as grammatical suffix prediction modulated by the left IFG. In speech processing, these cortical areas have the potential to support the function of stem tones as predictors for upcoming suffixes even at fast speech rates, where the suffix may well appear as early as 150 ms following tone onset.

Acknowledgments

This work was supported by the Swedish Research Council (Grant Number 2011-2284), the Crafoord Foundation (Grant Number 2011-0624), Knut and Alice Wallenberg Foundation (Grant Number 2014.0139), and Marcus and Amalia Wallenberg Foundation (Grant Number 2014.0039).

Appendix A. Supplementary material

Supplementary data associated with this article can be found, in the online version, at <http://dx.doi.org/10.1016/j.bandl.2015.07.009>.

References

- Ashburner, J., & Friston, K. J. (1999). Nonlinear spatial normalization using basis functions. *Human Brain Mapping*, 7(4), 254–266.
- Bland, J. M., & Altman, D. G. (1996). Calculating correlation coefficients with repeated observations: Part 2—Correlation between subjects. *BMJ*, 310, 633.
- Boudelaa, S., Pulvermüller, F., Hauk, O., Shtyrov, Y., & Marslen-Wilson, W. (2009). Arabic morphology in the neural language system. *Journal of Cognitive Neuroscience*, 22(5), 998–1010.
- Bruce, G. (1977). *Swedish word accents in sentence perspective*. Lund: Gleerups.
- Brunelliere, A., Hoen, M., & Dominey, P. F. (2005). ERP correlates of lexical analysis: N280 reflects processing complexity rather than category or frequency effects. *Neuroreport*, 16(13), 1435–1438.
- Chao, Y. R. (1976). *Aspects of Chinese sociolinguistics*. Stanford: Stanford University Press.
- Cutler, A., Dahan, D., & Donselaar, W. V. (1997). Prosody in the comprehension of spoken language: A literature review. *Language and Speech*, 40(2), 141–201.
- Dale, A. M. (1999). Optimal experimental design for event-related fMRI. *Human Brain Mapping*, 8, 109–114.
- Eng, N., Obler, L. K., Harris, K. S., & Abramson, A. S. (1996). Tone perception deficits in Chinese-speaking Broca's aphasics. *Aphasiology*, 10(6), 649–656.
- Friederici, A. D., & Alter, K. (2004). Lateralization of auditory language functions: A dynamic dual pathway model. *Brain and Language*, 89, 267–276.
- Friston, K. J., Ashburner, J., Frith, C., Poline, J. B., Heather, J. D., & Frackowiak, R. S. J. (1995). Spatial registration and normalization of images. *Human Brain Mapping*, 2, 165–189.
- Gandour, J., Dziedzic, M., Wong, D., Lowe, M., Tong, Y., Hsieh, L., et al. (2003). Temporal integration of speech prosody is shaped by language experience: An fMRI study. *Brain and Language*, 84, 318–336.
- Gandour, J., Jonglorpisit, S., Khunadorn, F., Dechongkit, S., Boongird, P., Boonklam, R., et al. (1992). Lexical tones in Thai after unilateral brain damage. *Brain and Language*, 43(2), 275–307.
- Gandour, J., Tong, Y., Wong, D., Talavage, T., Dziedzic, M., Xu, Y., et al. (2004). Hemispheric roles in the perception of speech prosody. *NeuroImage*, 23(1), 344–357.
- Gandour, J., Wong, D., Hsieh, L., Weinzapfel, B., Van Lancker, D., & Hutchins, G. D. (2000). A PET cross-linguistic study of tone perception. *Journal of Cognitive Neuroscience*, 12, 207–222.
- Graves, W. W., Grabowski, T. J., Mehta, S., & Gupta, P. (2008). The left posterior temporal gyrus participates specifically in accessing lexical phonology. *Journal of Cognitive Neuroscience*, 20(9), 1698–1710.
- Hirsch, J., Moreno, D. R., & Kim, K. H. (2001). Interconnected large-scale systems for three fundamental cognitive tasks revealed by functional MRI. *Journal of Cognitive Neuroscience*, 13(3), 389–405.
- Hughes, C. P., Chan, J. L., & Su, M. S. (1983). Aprosodia in Chinese patients with right cerebral hemisphere lesions. *Archives of Neurology*, 40, 732–736.
- Jung, T.-P., Makeig, S., Humphries, C., Lee, T.-W., McKeown, M. J., Iragui, V., et al. (2000). Removing electroencephalographic artifacts by blind source separation. *Psychophysiology*, 37, 163–178.
- Kaan, E., Barkley, C. M., Bao, M., & Wayland, R. (2008). Thai lexical tone perception in native speakers of Thai, English and Mandarin Chinese: An event-related potentials training study. *BMC Neuroscience*, 9(53).
- Kadymusuma, M. R., De Bleser, R., & Mayer, J. (2011). Perceptual discrimination of Shona lexical tones and low-pass filtered speech by left and right hemisphere damaged patients. *Aphasiology*, 25(5), 576–592.
- Klein, D., Zatorre, R. J., Milner, B., & Zhao, V. (2001). A cross-linguistic PET study of tone perception in Mandarin Chinese and English speakers. *NeuroImage*, 13, 646–653.
- Koester, D., & Schiller, N. O. (2011). The functional neuroanatomy of morphology in language production. *NeuroImage*, 55, 732–741.
- Lancaster, J. L., Summerlin, J. L., Rainey, L., Freitas, C. S., & Fox, P. T. (1997). The Talairach Daemon, a database server for Talairach Atlas Labels. *NeuroImage*, 5(4), S633.
- Lehmann, D., & Skrandies, W. (1980). Reference-free identification of components of checkerboard-evoked multichannel potential fields. *Electroencephalography and Clinical Neurophysiology*, 48, 609–621.
- Levelt, W. J. M. (1989). *Speaking: From intention to articulation*. Cambridge, MA: The MIT Press.
- Maldjian, J. A., Laurienti, P. J., Kraft, R. A., & Burdette, J. H. (2003). An automated method for neuroanatomic and cytoarchitectonic atlas-based interrogation of fMRI data sets. *NeuroImage*, 19, 1233–1239.
- Marslen-Wilson, W., & Tyler, L. (2007). Morphology, language and the brain: The decompositional substrate for language comprehension. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 362(1481), 823–836.
- Moen, I. (1993). Functional lateralization of the perception of Norwegian word tones: Evidence from a dichotic listening experiment. *Brain and Language*, 44, 400–413.
- Moen, I., & Sundet, K. (1996). Production and perception of word tones (pitch accents) in patients with left and right hemisphere damage. *Brain and Language*, 53, 267–281.
- Naeser, M. A., & Chan, S. W.-C. (1980). Case study of a Chinese aphasic with the Boston Diagnostic Aphasia Exam. *Neuropsychologia*, 18, 389–410.
- Neville, H. J., Mills, D. L., & Lawson, D. S. (1992). Fractioning language: Different neural subsystems with different sensitive periods. *Cerebral Cortex*, 2, 244–258.
- Packard, J. L. (1986). Tone production deficits in nonfluent aphasic Chinese speech. *Brain and Language*, 29, 212–223.
- Piras, F., & Paola, M. (2009). Word and number reading in the brain: Evidence from a voxel-based lesion-symptom mapping study. *Neuropsychologia*, 47, 1944–1953.
- Pulvermüller, F., & Shtyrov, Y. (2006). Automatic processing of grammar in the human brain as revealed by the mismatch negativity. *NeuroImage*, 20, 159–172.
- Riad, T. (2012). Culminativity, stress and tone accent in Central Swedish. *Lingua*, 122(13), 1352–1379.
- Rischel, J. (1963). Morphemic tone and word tone in Eastern Norwegian. *Phonetica*, 10, 154–164.
- Ritter, P., & Villringer, A. (2006). Simultaneous EEG–fMRI. *Neuroscience and Biobehavioral Reviews*, 30, 823–838.
- Roll, M., Horne, M., & Lindgren, M. (2010). Word accents and morphology—ERPs of Swedish word processing. *Brain Research*, 1330, 114–123.
- Roll, M., Horne, M., & Lindgren, M. (2011). Activating without inhibiting: Left-edge boundary tones and syntactic processing. *Journal of Cognitive Neuroscience*, 23(5), 1170–1179.

- Roll, M., Söderström, P., & Horne, M. (2013). Word-stem tones cue suffixes in the brain. *Brain Research*, 1520, 116–120.
- Sammler, D., Kotz, S. A., Eckstein, K., Ott, D. V. M., & Friederici, A. D. (2010). Prosody meets syntax: The role of the corpus callosum. *Brain*, 133(9), 2643–2655.
- Shtyrov, Y., & Pulvermüller, F. (2002). Memory traces for inflectional affixes as shown by mismatch negativity. *European Journal of Neuroscience*, 15, 1085–1091.
- Söderström, P., Roll, M., & Horne, M. (2012). Processing morphologically conditioned word accents. *The Mental Lexicon*, 7(1), 77–89.
- Tyler, L. K., Marslen-Wilson, W. D., & Stamatakis, E. A. (2005). Differentiating lexical form, meaning and structure in the neural language system. *Proceedings of the National Academy of Sciences of the United States of America*, 102, 8375–8380.
- Wang, Y., Sereno, J. A., Jongman, A., & Hirsch, J. (2003). fMRI evidence for cortical modification during learning of Mandarin lexical tone. *Journal of Cognitive Neuroscience*, 15(7), 1019–1027.
- Wang, Y., Jongman, A., & Sereno, J. A. (2001). Dichotic perception of Mandarin tones by Chinese and American listeners. *Brain and Language*, 78, 332–348.
- Worsley, K. J., Poline, J. B., Vandal, A. C., & Friston, K. J. (1995). Tests for distributed, nonfocal brain activations. *Neuroimage*, 2, 183–194.
- Xu, Y., Gandour, J., Talavage, T., Wong, D., Dziedzic, M., Tong, Y., et al. (2006). Activation of the left planum temporale in pitch processing is shaped by language experience. *Human Brain Mapping*, 27, 173–183.
- Yiu, E. M.-L., & Fok, A. Y.-Y. (1995). Lexical tone disruption in Cantonese aphasic speakers. *Clinical Linguistics and Phonetics*, 9(1), 79–92.