AARHUS UNIVERSITY

Coversheet

This is the accepted manuscript (post-print version) of the article.

Contentwise, the post-print version is identical to the final published version, but there may be differences in typography and layout.

How to cite this publication (APA)

Please cite the final published version:

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](https://doi.org/10.1016/j.neuroimage.2017.03.066)

Publication metadata

Title:	Flexible, rapid and automatic neocortical word form acquisition mechanism in children as revealed by neuromagnetic brain response dynamics
Author(s):	Partanen, E., Leminen, A., de Paoli, S., Bundgaard, A., Kingo, O. S., Krøjgaard, P., & Shtyrov, Y.
Journal:	NeuroImage
DOI/Link:	http://dx.doi.org/10.1016/j.neuroimage.2017.03.066
Document version:	Accepted manuscript (post-print)

General Rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognize and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Flexible, rapid and automatic neocortical word form acquisition mechanism in children as revealed by neuromagnetic brain response dynamics

Eino Partanen^{1,2}, Alina Leminen^{1,2}, Stine de Paoli¹, Anette Bundgaard¹, Osman Skjold Kingo³, Peter Krøjgaard³, and Yury Shtyrov^{1,4,5}

¹Center of Functionally Integrative Neuroscience (CFIN), Department of Clinical Medicine, Aarhus University, Aarhus, Denmark

²Cognitive Brain Research Unit, Institute of Behavioural Sciences, University of Helsinki, Finland

³Center on Autobiographical Memory Research (CON AMORE), Department of Psychology and Behavioural Sciences, Aarhus University

⁴Laboratory of Behavioural Neurodynamics, Saint Petersburg State University, Saint Petersburg, Russian Federation

⁵Medical Research Council (MRC), Cognition and Brain Sciences Unit, Cambridge, UK

Corresponding author:

Eino Partanen

eino.partanen@helsinki.fi

+358-40-5604141

Keywords: MEG, children, language, learning, development, ERF

Abstract

Children learn new words and word forms with ease, often acquiring a new word after very few repetitions. Recent neurophysiological research on word form acquisition in adults indicates that novel words can be acquired within minutes of repetitive exposure to them, regardless of the individual's focused attention on the speech input. Although it is well-known that children surpass adults in language acquisition, the developmental aspects of such rapid and automatic neural acquisition mechanisms remain unexplored. To address this open question, we used magnetoencephalography (MEG) to scrutinise brain dynamics elicited by spoken words and word-like sounds in healthy monolingual (Danish) children throughout a 20-minute repetitive passive exposure session. We found rapid neural dynamics manifested as an enhancement of early (~100 ms) brain activity over the short exposure session, with distinct spatiotemporal patterns for different novel sounds. For novel Danish word forms, signs of such enhancement were seen in the left temporal regions only, suggesting reliance on pre-existing language circuits for acquisition of novel word forms with native phonology. In contrast, exposure both to novel word forms with non-native phonology and to novel non-speech sounds led to activity enhancement in both left and right hemispheres, suggesting that more wide-spread cortical networks contribute to the build-up of memory traces for non-native and non-speech sounds. Similar studies in adults have previously reported more sluggish (~15-25 minutes, as opposed to 4 minutes in the present study) or non-existent neural dynamics for non-native sound acquisition, which might be indicative of a higher degree of plasticity in the children's brain. Overall, the results indicate a rapid and highly plastic mechanism for a dynamic build-up of memory traces for novel acoustic information in the children's brain that operates automatically and recruits bilateral temporal cortical circuits.

Highlights

- Children acquire novel word forms in an automatic fashion after mere 4 minutes.

- To acquire new native-language word forms, the left perisylvian neural network is used.
- Bilateral hemispheric structures support acquisition of non-native language word forms.

1. Introduction

By early adulthood, humans have acquired vocabularies of between 15,000 and 20,000 words (D'Anna et al., 1991), enabling efficient and accurate linguistic communication. However, to obtain such an extensive vocabulary, word learning in childhood has to be extremely rapid. Indeed, behavioural studies indicate that children learn words with apparent ease and new word forms are added into active vocabulary after only a handful of repetitions (e.g. Dollaghan, 1985). Although estimates for the rate of lexical acquisition in childhood vary depending on various factors, for example, the meaning of the new words (e.g. Goldfield & Reznick, 1990), some studies suggest the learning rate to be as high as 10-20 words per week (e.g. Ganger & Brent, 2004).

How are new words learned? The mainstream views maintain that consolidation of new word form representations and their integration into the mental lexicon is a lengthy process, which is preceded by a initial encoding stage that appears to be a swift and largely automatised process (e.g. Kimppa et al., 2015), with consolidation requiring at least an overnight sleep period (Gaskell & Dumay, 2003; Davis & Gaskell 2009). This initial process of rapid acquisition of a novel word form seems to be highly important for language learning, as behavioural studies have indicated that good performance in such tasks predicts improved language proficiency in children (e.g. Gathercole, 2006). However, despite the fact that rapid acquisition of novel word forms has been studied behaviourally for decades (see, e.g., Carey & Bartlett, 1978), the neurobiological foundations of both children's and adults' rapid learning abilities are poorly understood. Consequently, a recent line of research has strived to illuminate the neurophysiology of online word form acquisition (for a review, see Shtyrov, 2012). Indeed, recent studies in adults have shown that formation of neural memory traces for novel spoken word forms with native phonology takes place after mere 15-30 minutes of passive exposure to them (e.g. Shtyrov

et al., 2010; Yue, Bastiaanse, & Alter, 2014; Kimppa et al., 2015). In these studies, lexical memory trace formation is reflected as an early (~50-150 ms after recognition point) amplitude increase in the brain's event-related responses (ERPs) as a result of exposure to novel word forms. In adults, this enhancement takes place regardless of the individual's focused attention on the speech input and appears to be specific to novel word forms with native phonology only (e.g. Kimppa et al., 2015; see also Weber-Fox & Neville, 1996). Furthermore, this enhanced neural responsiveness as a result of exposure to novel word forms has been clearly associated with behavioural learning outcomes and individual learning profiles (Kimppa et al., 2015, 2016), supporting the notion that such activity is a genuine neural correlate of the acquisition process. Given that the neural changes reflecting acquisition of novel word forms in adults are very rapid, and given that children acquire new words in an extremely short time span (Dollaghan, 1985; Carey & Bartlett, 1978), it seems plausible to assume that the neural correlates of novel word form acquisition in children may manifest as adult-like – or even more efficient – plastic changes in brain activity. However, to the best of our knowledge, this hypothesised speed of brain activity change in children has never been tested empirically.

Even though it is plausible to expect children to exhibit neural processes reflecting memory trace formation for novel word forms comparable to adults, there is ample evidence to suggest that children's brain is much more malleable than that of adults. Thus, while neural processes underlying rapid and automatic word form acquisition in adults have revealed no evidence for rapid acquisition of novel word forms with non-native phonology (Kimppa et al., 2015), children's innate ability to acquire a new language from scratch is superior to that of adults (for a review, see Birdsong, 2006), at least prior to puberty (Johnson & Newport, 1989). Therefore, one might expect that this superior ability of children to acquire a new language may also be reflected as rapid formation of new neural memory traces for items with non-native phonology as well.

Similarly to non-native word forms, the question of rapid memory trace build-up can be asked about non-speech sounds as well. Previous research indicated that acoustically matched complex sounds do not elicit the same fast plastic changes in adult neocortex as word forms in native speech (Shtyrov, 2011). However, children's ability to acquire new word forms may differ from that of adults and could possibly be at least partially based on a more general learning mechanism. Thus, an open question remains as to whether children's brain is capable of rapidly forming novel representations for any acoustic events, or whether this phenomenon is restricted to language only.

On the neural sources of rapid acquisition of novel word forms reflected in the memory-trace related enhancement of neural activation, regions in the left perisylvian cortices, including left temporal and inferior frontal regions, seem to be most frequently activated when examining adults. However, it is unclear whether children exhibit activation in the similar brain regions when processing speech. While a recent meta-analysis comprising of fMRI studies on language processing in children between 0-18 years of age reported that adults and children utilise similar cortical regions, differences in activation patterns between, for example, low-level sensory regions and higher-level control regions, may change with age (Weiss-Croft & Baldeweg, 2015). These age-related differences between both children of different ages and between children and adults are consistent with the Skill Learning Hypothesis (Johnson, 2001) postulating that the activity in fronto-parietal regions, responsible for, e.g., top-down processing and forming a part of the attention network (Posner & Rothbart, 2007; Ptak 2012), diminishes with increasing age. In contrast, an opposite effect of activity increase is seen in temporal sensory regions, associated with the increased automatising of speech processing and categorisation. Thus, the present evidence from fMRI studies on language processing in childhood suggests that, at least for neural processes involving native language processing, similar cortical regions should be activated in children and adults. However, when a child's brain is acquiring novel items that include sounds that do not exist in her native language, the developing brain might still

need to recruit different cortical regions in the acquisition process. For instance, it may involve higher-order fronto-parietal regions, reflecting the degree of resource-intensive top-down processing required to process the incoming speech signal (Brauer & Friederici, 2007) in order to enable formation of a new memory trace in the brain's mental lexicon. Thus, a question arises on which neural network children's ability to acquire word forms with non-native phonology may rely.

Adult studies on processing of non-native versus native speech have shown similar brain regions being activated, i.e. mainly perisylvian regions (Wilson & Iacoboni, 2006; Golestani & Zatorre, 2004). Some studies suggest, however, that the areas utilised for speech processing may depend on the cues that are required to identify the speech sounds (Zhao et al., 2008), suggesting different activation patterns for native and non-native speech. Furthermore, studies investigating learning or acquisition of novel word forms have shown that even right-hemispheric perisylvian structures are involved in the acquisition of novel items (Paulesu et al., 2009), which may reflect either genuine lexicalisation or lexical retrieval of specific items (Damasio et al., 1996), or additional attentional or working memory processes required during stimulus processing (Vigneau et al., 2011). Thus, partially different neural structures might be utilised during acquisition of novel words with native versus non-native phonology, possibly due to the different acoustic cues required to process these words and different degrees of top-down processing of non-native sounds. However, in the absence of experimental evidence, only tentative hypotheses can be made.

Given the lack of research on the neural processes underlying novel word form acquisition in children, we designed the current experiment to fill these gaps. To assess the automatic word form acquisition, we presented 5-12 year old children with real words, novel word forms incorporating either native or non-native phonology, and equally complex non-speech control sounds. Non-speech sounds mimic the complex acoustic properties of a speech signal but cannot be interpreted as human speech. To accurately capture the dynamics associated with online exposure to novel word forms, we used

magnetoencephalography (MEG), a neuroimaging method optimally combining spatial and temporal resolution, allowing us to specify the neural memory trace formation process in both space and time. In our analyses, we scrutinised the brain dynamics elicited by these sounds throughout the 20-minute passive exposure session. Cortical generators of neural activity underlying surface MEG dynamics were analysed using distributed source reconstruction techniques (weighted minimum-norm current estimates). In accordance with the results previously obtained with adults, we expected to find changes in neural dynamics, reflecting putative formation of neural memory traces for novel word forms, in left perisylvian regions. Regarding exposure to new items with non-native phonology, we expected that partially different neural networks would be recruited during acquisition of these items, possibly reflecting the increased need for top-down processing of such items and involving other sources, including those in the right hemisphere.

2. Material and methods

2.1. Participants

Twenty three monolingual native Danish-speaking children successfully participated in the MEG experiment (12 male; mean age 9.1 years, range 5.9 – 12.75 years). A larger age range was chosen as to control for whether word form acquisition was associated with reading development, which would be superior in older relative to younger children. None of the parents of the participants reported their children to have any neurological deficits. The study was evaluated by the Central Jutland Regional Ethical Board and conducted in accordance with the Helsinki Declaration. Written consent was obtained from both the children and their parents; if a child was unable to read, oral consent was obtained instead.

To control whether neural word form acquisition processes were associated with general cognitive functions or reading ability, several performance tests were conducted. Performance IQ (PRI), verbal

IQ (VCI) and working memory index (WMI) of the children were assessed using WISC-IV test (Wechsler, 2004). The PRI was evaluated with a block design subtest, the VCI with the verbal comprehension subtest while digit span subtest was used to gauge WMI. The mean VCI of the children was 9.61 (SD 1.64), PRI 15.43 (SD 2.57), and WMI 9.22 (SD 1.86). Finally, reading ability was tested by asking children to read aloud a short (45 words) story while a researcher marked down any mistakes the child made.

2.2. *Stimuli*

We created a balanced stimulus set comprising acoustically and psycho-linguistically controlled real words, phonologically native novel word forms, novel word forms with non-native phonology and complex non-speech sounds. In order to minimise the acoustic variance between the stimuli, two initial consonant-vowel (CV) fragments were generated: ka ([ka]) and ko ([kʌ]) and matched for length (270 ms) and loudness (measured as root-mean-square [RMS] power). These two fragments were then cross-spliced with final consonants t ([d]), k ([g]), also matched for length and RMS loudness, to generate either two real Danish words (kat; ['kad], 'cat' – log surface frequency 7.12, and kok; ['kʌg], 'cook' – log surface frequency 8.01 (KorpusDK, Retrieved May 30, 2016)), or two novel Danish word forms (kak; ['kag] and kot; ['kʌd]). Furthermore, two novel word forms with non-native phonology containing a French nasal vowels (['kãg] and ['kõd]) were generated by cross-splicing the initial CV-fragments ['kã] and ['kõ], matched for length and RMS loudness with the Danish consonant-vowel fragments, with the two final consonants used with the real and novel Danish word forms, t ([d]), k ([g]). All sounds were spoken by the same female native Danish/French bilingual.

Finally, two so called “musical rain” sounds, which share the complex auditory properties of a speech signal but cannot be interpreted as speech (see, e.g., Uppenkamp, Johnsrude, Norris, Marslen-Wilson, & Patterson, 2006, Figure 1, for graphical illustration of the stimuli generation process), were generated from the two aforementioned Danish words. These musical rain stimuli were generated as

follows. First, the formant time course of the Danish words 'kat' and 'kok' were extracted using Praat software (version 5.4.01; Boersma, 2002). Second, their pitch was randomly shifted up or down up to half of an octave and the envelope onset of the formant was randomised. Third, the final waveform was generated by adding the time courses of these formants together, and a band-pass filter between 20 and 24000 Hz was applied to remove any possible high or low frequency artifacts. These non-speech sounds were included to assess whether the rapid neural dynamics underlying word form acquisition were specific to speech only, and if not, to examine whether the neural sources underpinning the enhancement due to exposure to these non-speech sounds would differ from those of speech sounds.

Thus, the disambiguation point between spoken Danish word forms, non-native word forms and musical rain sounds was at the stimulus onset, while the native words and novel word forms only differed after 270 ms from stimulus onset. The stimuli were validated in a perceptual rating experiment, in which three adult native speakers of Danish were presented the Danish real words and the novel word forms and asked to write down what they heard; a 100% correct identification rate was obtained.

2.3. Procedure

During the experiment the children were seated in an electromagnetically and acoustically isolated MEG chamber. They were presented with all eight stimuli repeated in a pseudo-random order (no more than two subsequent repetitions of the same stimulus) for 20 minutes, during which they were watching a self-selected silent film. The children were instructed not to pay any attention to the sounds and to focus on the film instead. A total of 1200 stimuli were presented (i.e. 150 repetitions of each stimulus with $p=1/8$) with 1 second stimulus onset asynchrony (SOA). While the children were in

the shielded room, they were always accompanied by a researcher, or if the child so wished, a parent, who instructed the children to focus on the film if their attention lapsed.

2.4. MEG recording and data analysis

Continuous MEG was recorded using an Elekta Triux™ whole-head 306-channel MEG system (ElektaNeuromag®, Elekta Oy, Helsinki, Finland) with 1000 Hz sampling rate and pass band from 0.03 to 200 Hz. Electrodes were placed at the temples to monitor the horizontal electro-oculogram (EOG), and above and below the right eye to monitor vertical EOG. The head position inside the MEG device was assessed by tracking locations of four head position indicator (HPI) coils in relation to the cardinal points of the head (nasion, left and right preauricular points), which were identified prior to the experiment with an Isotrak 3D-digitizer (Polhemus, Colchester, VT, USA). To remove externally generated magnetic artifacts and correct for head movements during the recording session, the continuous MEG raw data were pre-processed offline using a spatiotemporal signal space separation (tSSS) method as implemented in the MaxFilter™ software (ElektaNeuromag®, Elekta Oy, Helsinki, Finland). Data were analysed further using Brainstorm (Tadel et al., 2011) software. Eye blinks were detected on the basis of the activity in the bipolar vertical EOG channel, and corrected for using Brainstorm's signal-space projection (SSP) algorithm, as were artifacts arising from heart beats. After SSP processing, the data were bandpass filtered (0.1 to 45 Hz) in accordance with previous MEG studies (Leminen et al., 2011) and in order to exclude the noise arising from continuous high positioning process, and divided into epochs of 1000 ms, starting from 100 ms prior to stimulus onset. All epochs were baseline corrected to pre-stimulus baseline from -100 to 0 ms, following which all epochs where MEG activity exceeded 4000 fT/cm on any gradiometer channel were removed from the analysis.

To assess the temporal dynamics of neural activity associated with exposure, individual participant averages were formed as follows. First, the data were divided into five equal bins according to the duration of exposure to the stimuli, with each bin containing 20% of the experimental recording (i.e. ~4 minutes). Next, event-related responses for the four different stimulus categories were formed separately for each bin, participant and stimulus category (real words, novel native word forms, novel non-native word forms, musical rain). Finally, group-average ERFs were formed by averaging across the participants for each bin and stimulus category. The mean number of accepted epochs in each group-average was 44 (median 45, range 19 – 57). In order to determine the most prominent responses in the event-related field (ERF) waveform in an unbiased data-driven fashion, a group-average global field power waveform was formed from all the accepted epochs and gradiometer pairs averaged across all speech sounds used in the experiment (Figure 1). On the basis of visual inspection, three peaks were identified: 91, 317, and 452 ms from stimulus onset. As in previous studies (Shtyrov et al., 2010, Yue, Bastiaanse & Alter, 2014; Kimppa et al., 2015) showing rapid lexical acquisition effects at latencies around ~50-120 ms, we a priori focussed on the first two peaks as they fall within 200 ms after the disambiguation points (see Figure 1, and the Stimuli section above). Ad hoc, we also analysed the peak at 452ms as well to verify that no similar effects were present at a later stage.

Figure 1. Global field power waveform across all the accepted epochs and gradiometer pairs for all speech sounds used in the experiment. Three peaks were identified from the waveform, at 91, 317, and 452 ms from stimulus onset, respectively. Stimulus waveform examples are displayed overlaid on the GFP waveform. Familiar and unfamiliar native word forms differ from each other only at approximately 270 ms from stimulus onset (divergence point marked by magenta line). In contrast, the disambiguation point between native and non-native word forms and non-speech sounds (bottom) is at stimulus onset (cyan-grey line).

We assessed the temporal dynamics of neural activity associated with repetitive exposure to novel items by analysing relative amplitude changes between first time bin (first 20% of the experiment) and the remaining time bins of the experiment, in keeping with previous studies scrutinising online ERP/ERF dynamics. To that end, data from each individual participant were normalised relative to

the response magnitude in the first bin (*relative response amplitude* = $\frac{\text{response magnitude in bin } n}{\text{response magnitude in bin } 1}$), calculated separately for each individual participant, stimulus category, and time point). Finally, initial analysis indicated that the changes in the response dynamics were confined to the boundary between the first and second time bin and the responses remained stable for the rest of the experiment (see Supplementary data 1 and 2). Thus, data were further averaged into initial (first time bin) and final (bins 2-5) segments, separately for each stimulus category and participant. For these final averages, the mean number of accepted epochs was 38 (75% of all epochs; 36 epochs or 72% for familiar word form, 38 epochs or 77% for unfamiliar word forms, 39 epochs or 77% for nonnative word forms, and 38 epochs or 77% for non-speech sounds) for the initial segment and 174 for the final segment (73% of all epochs; 175 epochs or 73% for familiar word form, 173 epochs or 72% for unfamiliar word forms, 175 epochs or 73% for nonnative word forms, and 176 epochs or 73% for non-speech sounds; all differences n.s.). The mean ERF amplitudes in 40 ms windows centred at 91 and 317 ms from stimulus onset were determined from each participant from a cluster of gradiometer pairs over the temporo-frontal perisylvian areas, separately for left and right hemispheres. The criteria to include a channel in the cluster was based on visual inspection, individually for each child: if the gradiometer pair waveform showed identifiable peaks clearly above the baseline noise level at the same latencies as the individual GFP waveform of all sensors, it was included in the cluster. This approach was necessary as childrens' headsizes varied widely and many children tilted their head downwards to variable degree during the MEG recording, , reducing the head surface coverage by the MEG array and resulting in varying number os channels with identifiable neural responses. Thereby defined clusters comprised on average 9 gradiometer pairs (18 gradiometers). To assess differences between initial and final segments of the experiment, two-tailed t-tests were run, separately for both hemispheres. To evaluate effect sizes, Cohen's *d* was used.

Where a statistically significant relative difference between initial and final segments were found in the ERF analyses, underlying cortical generators of the respective ERF effects were estimated using distributed source reconstruction techniques (weighted minimum-norm current estimates, as implemented in Brainstorm software; see e.g. Lin, Witzel, Ahlfors, Stufflebeam, Belliveau, et al., 2006). Age-appropriate MRI templates, with average templates available for each age group with half year intervals (e.g. 6 years, 6.5 years, 7 years), obtained from the Neurodevelopmental MRI Database (Evans & Brain Development Cooperative Group, 2006; Sanchez & Almli, 2012; Richards & Xie 2015; Richards, Sanchez, Phillips-Meek, & Xie, 2016), were segmented using Brainsuite (version 14c; <http://brainsuite.org/>) software and using the BCI-DNI brain atlas (Pantazis et al., 2009) , which were then utilised to create overlapping spheres head models (Huang et al., 1999) in Brainstorm software. Based on the previous studies on neuroanatomical substrates of language (e.g. Hickok & Poeppel, 2007), source activity waveforms were extracted from regions of interest (ROIs) defined a priori and consisting of established language areas in the temporal lobe, temporal pole, inferior frontal, and supramarginal regions. Similar to the ERF analysis above, the extracted neural activity was normalised relative to the response magnitude in the first bin and statistically assessed using one-tailed t-tests corrected for multiple comparisons by the number of ROIs using the FDR correction (both corrected and uncorrected values are reported). To assess differences between different stimulus types, an ANOVA with Stimulus (familiar word form, unfamiliar word form, nonnative word form, non-speech sound) and Hemisphere (left, right) as within-subjects factors were conducted in sensor space while an ANOVA with Stimulus (4), Hemisphere (2) and ROI (temporal lobe, temporal pole, inferior frontal, and supramarginal regions) as within-subjects factors were conducted for source level data, Greenhouse-Geisser corrected for sphericity violations when appropriate and partial η^2 s reported for effect sizes. Finally, Spearman's correlation analyses were used to investigate the effects of background factors (age, VIQ, PIQ, WMI, reading fluency) on both neural and source activity (rates of change from initial to final segments).

3. Results

In the ERF analyses statistically significant differences between initial and final parts of the experiment were found for all novel stimuli used in the experiment. For novel word forms with non-native phonology, such a difference was observed in the right hemisphere at 91 ms from stimulus onset ($t(22)=2.237$, $p<0.036$, Cohen's $d=0.95$) whereas comparable effects for musical rain sounds at 91 ms from stimulus onset manifested on both left ($t(22)=2.244$, $p<0.035$, Cohen's $d=0.96$) and right hemispheres ($t(22)=2.134$, $p<0.044$, Cohen's $d=0.91$). In a similar vein, enhancement of neural activity for native novel word forms was evident at 317 ms from stimulus onset (i.e. 47 ms from the uniqueness point) in the left hemisphere only ($t(22)=2.215$, $p<0.037$, Cohen's $d=0.94$). For familiar native word forms, the enhancement did not reach significance ($t(22)=1.652$, $p>0.112$, Cohen's $d=0.34$); see Figure 2 for illustration. For the peak observed at 452 ms, no statistically significant results were found. Analysing differences in activity enhancement between different conditions in sensor space resulted in a near-significant interaction ($F(3,66)=2.766$, $p<0.069$, $\eta^2=0.112$). Post-hoc tests indicated this to be due to response enhancement to nonnative word forms and because non-speech sounds tended to be larger on the right hemisphere than response enhancements to unfamiliar native word forms ($t(22)=1.991$, $p<0.059$, $d=0.85$ and $t(22)=1.823$, $p<0.082$, $d=0.77$ respectively).

Figure 2. Panel A depicts ERF waveforms (using absolute values; for illustration-n only) from the cluster of gradiometer pairs for native novel word forms (magenta), native familiar words (blue), novel word forms with non-native phonology (cyan) and musical rain sounds (grey). For native novel and familiar word forms, ERF waveforms display the cluster of sensors on the left hemisphere, whereas right-hemispheric ERF waveforms depict non-native novel word forms and musical rain sounds. Vertical dotted lines denote when the disambiguation point from the familiar native word form and gray vertical rectangles denote the first peak of interest after the disambiguation point. Panel B, on right, shows relative changes between initial (solid lines) and final (dashed lines) parts of the experiment. For native novel word forms, significant differences were found in the left hemisphere only; for non-native novel word forms, effects were statistically significant in the right hemisphere; for musical rain sounds, significant differences between initial and final parts of the experiment were found over both hemispheres. Panel B, on left, illustrates the relative ERF topographies of relative changes in at peak latencies: *: $p < 0.05$

Analysing the neural sources underlying the observed response enhancement, a number of statistically significant effects in core language regions were found (see Figure 3 for a summary). Specifically, for the word forms with non-native phonology (top panel, cyan), source activations were found to be significantly increased with exposure in the left supramarginal gyrus ($t(22)=2.317$, $p < 0.044$ corrected, Cohen's $d=0.48$; $p < 0.015$ uncorrected), left temporal region ($t(22)=2.2056$, $p < 0.044$ corr; Cohen's $d=0.43$; $p < 0.026$ uncorr), and right temporal pole ($t(22)=2.368$, $p < 0.044$ corr, Cohen's $d=0.35$; $p < 0.0135$ uncorr). For the musical rain sounds (top panel, gray), effects pertaining to response enhancement were seen in the right supramarginal gyrus ($t(22)=3.043$, $p < 0.009$ corr, Cohen's $d=0.63$; $p < 0.003$ uncorr) and right temporal region ($t(22)=3.213$, $p < 0.009$ corr, Cohen's $d=0.67$; $p < 0.002$ uncorr); further effects were localised to the left temporal pole ($t(22)=2.971$, $p < 0.009$ corr, Cohen's $d=0.62$; $p < 0.004$ uncorr). For the phonologically native novel words (bottom panel, magenta; real

words in blue), the statistical trend suggesting word form acquisition effects were seen in the left supramarginal gyrus ($t(22)=2.056$, $p<0.073$ corr, Cohen's $d=0.53$; $p<0.009$ uncorr) and left temporal region ($t(22)=2.089$, $p<0.096$, Cohen's $d=0.44$; $p<0.024$ uncorr).

Assessing differences in activity enhancement on source level between conditions, near-significant main effect of Stimulus ($F(3,66)=2.661$, $p<0.064$, $\eta^2=0.108$), Hemisphere ($F(1,22)=3.394$, $p<0.08$, $\eta^2=0.134$, suggesting slightly greater response enhancements in the left than right hemisphere) and a marginally significant interaction between Stimulus, Hemisphere, and ROI ($F(9,198)=2.081$, $p<0.055$, $\eta^2=0.86$) emerged. Post-hoc tests indicated that the response enhancement for familiar word forms was smaller than that of nonnative word forms or non-speech sounds ($t(22)=2.352$, $p<0.028$, $d=1.00$ for both). Post-hoc tests on the Stimulus x Hemisphere x ROI -interaction revealed that in the left supramarginal gyrus, the response enhancement for unfamiliar and nonnative word forms seemed greater than that of familiar words ($t(22)=2.670$, $p<0.014$, $d=1.14$ and $t(22)=2.054$, $p<0.052$, $d=0.88$ respectively) whereas in the right supramarginal gyrus, the response enhancement of non-speech sounds exceeded that of familiar ($t(22)=2.974$, $p<0.007$, $d=1.27$) and nonnative word forms ($t(22)=2.556$, $p<0.018$, $d=1.08$). In the right temporal pole, response enhancements to nonnative word forms were greater than those to unfamiliar word forms or non-speech sounds ($t(22)=2.137$, $p<0.044$, $d=0.91$ and $t(22)=2.387$, $p<0.026$, $d=1.02$, respectively). Finally, in the right temporal lobe, nonnative word forms and non-speech sounds elicited response enhancements of larger magnitude than those of words ($t(22)=3.041$, $p<0.006$, $d=1.30$) or non-speech sounds ($t(22)=3.792$, $p<0.001$, $d=1.62$). Response enhancement to non-speech sounds in the right temporal lobe was also stronger than that of unfamiliar word forms ($t(22)=2.865$, $p<0.009$, $d=1.22$). See Figure 3.

No statistically significant effects were found when correlating the relative changes in the ERF amplitudes or the source activations with any of the background variables (age, VIQ, PIQ, WMI, reading fluency; see supplementary data 3 for details).

Figure 3. Relative exposure-related changes in the source activity amplitudes for familiar word forms (blue), native novel word forms (magenta), non-native word forms (cyan) and musical rain sounds (grey) are shown as bars (error bars denote standard error of the mean). TP: temporal pole; TL: temporal region; SMG: supramarginal region (all values FDR corrected). Statistically significant differences between conditions are indicated using horizontal lines. #: $p < 0.1$; *: $p < 0.05$; **: $p < 0.01$

4. Discussion

Here, we present evidence of putative neural processes underlying rapid word acquisition in children, previously studied using behavioural methods only. We used a method similar to previous studies in adults – passive repetitive exposure to novel items – and analysed relative changes in the brain's neuromagnetic responses to these sounds during the short exposure session. During the experiment we presented unknown word forms with either native or non-native phonology and non-speech items repetitively, and used acoustically similar but familiar real words as controls. Our results generally indicate that 5-12 year old children exhibit neural dynamics somewhat similar to that previously described for adults (as known from earlier studies using comparable paradigms, e.g. Kimppa et al., 2015; see also Weber-Fox & Neville, 1996): rapid enhancement in ERF amplitudes as early as ~50-100 ms after the uniqueness point, with possibly shorter latencies for native-like stimuli (~50 ms) than for non-native or non-speech ones (~90 ms; see Figure 1). This enhancement is likely associated with formation of memory traces for novel sounds, albeit with two important distinctions. First, the neural dynamics associated with memory trace formation for novel items appeared to take place much faster in children than previously demonstrated in adults (approximate 4 minutes in the present study, in comparison to ~15-25 minutes in adults). This, and any other differences from previous studies, may be due to plethora of reasons, such as differences in stimulus parameters between the present study and previous adult studies, and needs to be interpreted with caution (particularly as such between study differences cannot be verified statistically). Second, changes in the neural dynamics were elicited by *all* novel stimuli, and not just phonologically native ones as seen in adults (Kimppa et al., 2015), although with diverging topographies. We tentatively hypothesise that this might be due to children's brains being more flexible in terms of memory trace formation for novel items as compared to adults. For novel native (Danish) word forms, the source analyses were partially inconclusive (which is likely related to difficulties related to source reconstruction in brains much smaller than the MEG

dewar and excessive motion artifacts in children) but still suggested that the enhancement was underpinned by a left-hemispheric perisylvian network, possibly indicating reliance on pre-existing phonetic/phonological circuits for acquiring words with native phonology. In contrast, exposure to new word forms with non-native phonology intensified the activity in both left and right hemispheric networks, suggesting that additional neural resources were recruited for the online acquisition of such items. Finally, the increase of neural activity was observed not only for speech but also for complex non-speech sounds. Unlike speech, this dynamic was bilateral, suggesting the use of distinct neural sources not utilised in acquisition of word forms. These findings are discussed further below.

4.1. Neural dynamics and substrate of online memory-trace build-up in children

The very rapid changes in neural dynamics, apparent ~50-100 ms from stimulus identification points, were seen as an enhanced neural activation following a short passive repetitive exposure to novel items. Typically, repetition of a stimulus leads to a reduction of brain responses, associated with both habituation and refractoriness (e.g. Budd et al., 1998) as well as with higher-order priming effects (e.g. Cabeza & Nyberg, 2000). In contrast, the activation seen in the present study for novel word forms with both native and non-native phonology indicates response enhancement which not only counteracts but overrides any repetition-related suppression. Although not directly comparable, these results tentatively support previous findings in adults (e.g., Shtyrov et al., 2010; Yue, Bastiaanse, & Alter, 2014; Kimppa et al., 2015, 2016): similarly to the findings with adults, we observed an increase of activation to novel word forms in the left temporal region, which may be indicative of adults and children utilising at least partially similar networks for rapid formation of memory traces for novel word forms. As previous studies indicate that the language lateralisation is fully established by the age of 5 years (or earlier), with only minimal changes later in development (Weiss-Croft & Baldeweg, 2015), this response enhancement in the left temporal region is highly indicative of reliance on a neural network tuned for native language processing.

In addition to the left temporal region, our results also indicate that enhanced activation of the left supramarginal region may be associated with word memory trace build-up in children. The left supramarginal region, and more generally the inferior parietal region, is thought to be involved in phonological processing and verbal working memory processes (e.g. Deschamps et al., 2014). It has also been reported to activate during pseudo-word processing, thus possibly reflecting attempts at lexical access (Newman & Twieg, 2001). In children, this region has been suggested to involve, for example, learning of sensorimotor patterns in speech (Shum et al., 2011). Indeed, in the developing brain, this region might serve two functions: first, to phonologically encode the novel word in working memory, and second, to incorporate the novel word into the pre-existing lexicon. Alternatively, if children's language processing genuinely relies more on top-down than automatised bottom-up processes (Weiss-Croft & Baldeweg, 2015), the activation in the left supramarginal region might also reflect changes in top-down regulation associated with memory trace formation for these novel word forms.

However, it pays to note that for the unfamiliar native language word forms, the source activation did not reach statistical significance when corrected for multiple comparisons. This may be caused by several factors, such as poor source resolution due to the use of adult size MEG or the rather large ROIs a priori employed in the present study (for safety reasons, individual children's MRIs were not obtained). It seems also plausible that for unfamiliar native language word forms, less neural processes may be required for their processing and incorporation into the brain's neural lexicon than for unfamiliar for non-native word forms (James & Gauthier, 2006), resulting in smaller neural activation which did not quite reach statistical significance in the present study. This phenomenon should be further investigated in future studies, with larger number of subjects, and, if only possible, using individual head models and smaller MEG dewars.

4.2. Speed and automaticity of memory-trace formation dynamics

An important difference between our study and the previous adult studies on rapid formation of neural memory traces for novel linguistic items is the dynamic changes in the brain responses. While acquisition-related changes in neural dynamics in adults are seen after approximately 15-25 minutes, in children, we observed equivalent changes already after 4 minutes of passive exposure, corresponding to only 30 repetitions of each individual word form. As a note of caution, we would like to note that this relative difference between adults and children may not be conclusively explained by the age difference alone and may originate from a number of methodological differences: the current study and the previous adult ones have employed different techniques (EEG vs. MEG), different number of stimuli for each category, different word complexities (disyllabic vs monosyllabic) and so forth. Future studies examining and directly comparing word form acquisition in adults and children would therefore presumably benefit considerably from using similar stimuli and paradigms. However, these confounds aside, this rapid rate of word form acquisition is not surprising in light of available behavioural evidence: in optimal settings children as young as two years can already acquire both the form and the meaning of a new word form even after a single exposure (e.g. Spiegel & Halberda, 2011). However, such previous behavioural studies have investigated children's language acquisition ability only in the context of an active attentive learning task. Moreover, these tasks made the meaning of the new word explicit (see, e.g., Gershkoff-Stowe & Hahn, 2007). Here, however, we show that the response increase to novel word forms takes place not only without explicitly providing the meaning of the word form, but even in absence of the children's attention. This suggests that the early stages of word acquisition can be at least partially automatic, an ability hitherto only demonstrated in adults (e.g. Kimppa et al., 2015; López-Barroso et al., 2016). Our results thus complement the previous behavioural findings by suggesting that children's rapid learning ability is not limited to active and attentive tasks. In fact, even a passive exposure to novel items automatically leads to extremely fast online changes in children's neural dynamics, suggesting a highly efficient memory trace build-up mechanism in the young brain. However, as we did not explicitly manipulate attention and only used

a 'non-attend' design, further studies are needed to assess the degree of attention (in)dependence of word form acquisition in children, which could be done by systematically manipulating the level of attention on the linguistic input.

4.3. Flexibility of novel memory-trace acquisition

The other striking result not previously observed in adults is that the children exhibited neural dynamics indicative of memory trace formation also for word forms with non-native phonology as well as for non-speech items. For the word forms with non-native phonology, the response increase was observed in both hemispheres. The enhancement of neural activation for word forms with non-native phonology in the right hemisphere requires an explanation. Previous studies have shown that the regions in the right hemisphere are involved in the processing of prosodic or pitch-related features of speech (for a review, see Lindell, 2006), but the right hemisphere may also be involved in language processing in more complex ways as well. For example, studies on second language (L2) acquisition suggest that bilinguals recruit resources in the right hemisphere for L2 processing (e.g. Dehaene et al., 1997; see also Gernsbacher & Kaschak, 2003, for a review), although this may depend on the age of L2 acquisition (Hull & Vaid, 2007). The right hemisphere's involvement was also shown to increase for native-language phonological processing, when stimuli were presented in acoustically unfavourable conditions, suggesting that right hemisphere activity increases when more neural resources are demanded (e.g. Shtyrov et al., 1999; Liikkanen et al., 2007). While the right-hemispheric activation seen for word forms with non-native phonology might reflect an increase of activation in a resource-demanding automatic categorisation task, such activation should still decrease over the course of the experiment due to repetition suppression effects (Cabeza & Nyberg, 2000), which is clearly not the case here. Thus, one might speculate that the changes in the dynamics in the right hemisphere may be associated with genuine neural memory trace formation. However, we are cautious in making this statement in absence of behavioural learning outcomes. Furthermore, another reason for caution is

the intriguing neural sources of this activation increase, localised at the right temporal pole. Whereas the left temporal pole has been implicated as a semantic hub in adults (Rogers et al., 2004), some lesion studies suggest that the right temporal pole is also involved in lexical retrieval, specifically naming (Damasio et al., 1996). Although the results from these lesion studies are in accordance with the novel lexical memory trace interpretation of our current data, the evidence is very sparse. In general, it seems that the apparent formation of new memory traces for items with non-native phonology require cannot be solely accomplished by the pre-existing neural circuits for native language phonemes but requires processes from other brain regions as well (see, e.g., Best & Tyler, 2007). However, future studies should further investigate the processes underlying non-native novel word learning in more detail, by linking the neural activity to behavioural outcomes by using, for example, auditory recognition and lexical decision tasks.

Even more intriguingly, our results also indicate an increase of neural responsiveness to non-speech “musical rain” sounds over the course of the experiment. This seems consistent with previous studies suggesting that auditory signatures of even meaningless or unpredictable sounds are automatically formed, possibly due to rapid sensory plasticity (Agus, Thorpe & Pressnitzer, 2010). It might be argued, then, that such process could be different from rapid acquisition of word forms and be subserved by different neural structures. Consistent with this, the changes in response dynamics in our study were observed both in the region including the primary auditory cortex (right temporal region), as well as in higher-level control regions, namely the left temporal pole and right supramarginal region. As sparse evidence obtained from adult studies suggests that processing of musical rain sounds is mainly managed by primarily basic auditory cortices, not higher-order auditory regions (e.g., Uppenkamp, Johnsrude, Norris, Marslen-Wilson, & Patterson, 2006), it may be that changes in dynamics observed in the present study are more indicative of stimulus acquisition functions than processing functions, or suggest differences in the processing of such stimuli between children and adults. Note, however,

that in previous studies the musical rain sounds were presented in isolation, not in context of other categorisable sounds. Furthermore, the children's processing of musical rain sounds may, to a larger degree, rely on higher-level control regions in the brain (Weiss-Croft & Baldeweg, 2015), which might also explain the differences between the adult studies and the present children data.

Nevertheless, as these sounds cannot be classified as speech and given that there should be no ready template to use for their discrimination, what could be the reason for this increase observed over both hemispheres? First, the fact that children are able to quickly form neural memory traces even for these uncategorisable sounds, suggests a great degree of flexibility in the child's brain, which seems to be capable of rapid memory trace formation of any novel acoustic event. This capacity might be crucial for the language learning ability in children. An alternative explanation is provided by the Accumulation Model framework, which reassesses the repetition suppression phenomena and proposes that both repetition suppression and repetition enhancement effects may be elicited simultaneously (James & Gauthier, 2006). In the Accumulation Model, stimulus processing is a continuous process with additional neural resources allocated to the recognition process until recognition is achieved (James & Gauthier, 2006), which may not be attained for the uncategorisable musical rain sounds. Thus, the bilateral enhanced neural activation for these non-speech sounds here might be indicative of the brain allocating more neural resources to categorise these sounds, somewhat similar to one of the interpretations put forward above for the non-native speech stimuli. Thus, assuming that the brain is incapable of categorising or forming neural memory traces for the musical rain sounds, then the enhanced activation in the left temporal region over the course of the experiment might be indicative of complex auditory processing. Further, the end result of these neural computations might be fed forward to the temporal pole, where activity increases as a result of longer-lasting and unsuccessful semantic search. Moreover, the observed enhanced activation in the right supramarginal gyrus might also be indicative of complex auditory processing, as this region has been

shown to play a role in, for example, recognition and processing of tones (Binder et al., 1997) or suprasegmental cues in speech (Geiser et al., 2008).

It should be noted that in the absence of behavioural learning outcomes of these sounds, we cannot confidently claim that the effects seen for musical rain stimuli are indicative of memory trace formation or mere complex auditory processing. However, the neural sources underlying the enhancement effects do seem to differ drastically between the musical rain stimuli and both the native and non-native word forms. Hence, this issue remains to be further addressed in future studies. A further topic of interest is the distinct lack of correlations between the individual rate of neural response change and age or performance measures (VIQ, PIQ, WMI, reading skill). It may be that the rapid acquisition mechanism develops before the age of 6 years and functions in somewhat similar fashion in neurotypical children of all ages; it would, therefore, be of interest to study this phenomenon in either younger children or in children with linguistic developmental deficit, such as SLI or dyslexia.

5. Conclusions

Our results illustrate, for the first time, the spatiotemporal dynamics of rapid and automatic build-up of neural memory traces for word forms with both native and non-native phonology in children. They suggest that the developing brain utilises a left-lateralised perisylvian network for acquiring novel word forms, involving at least the left temporal lobe and supramarginal regions. In comparison with previous studies (which is indirect and should thus be treated with extreme caution), partially similar regions are used in the word form acquisition process by both children and adults. However, the children's brains seem more malleable than those of adults in acquiring novel word forms, as changes in neural dynamics were observed for word forms with non-native phonology as well. It seems that in order to efficiently form neural memory traces for non-native word forms, children recruit additional

neural resources, underpinned by changes not only in the left- but in the right-hemispheric regions as well. Finally, we also found bilateral changes in response dynamics for non-speech items as well. We conclude that the developing brain has neural mechanisms in place enabling it to acquire novel items faster than known for the adult nervous system, and that is still sufficiently plastic to automatically acquire novel word forms with non-native phonology as well as non-speech items.

Acknowledgments

We would like to thank Asta Tybirk Le Marc for her invaluable help in recording the stimuli used in the experiment. This work was supported by the Lundbeck foundation, Denmark.

References

- Agus, T. R., Thorpe, S. J., & Pressnitzer, D. (2010). Rapid formation of robust auditory memories: insights from noise. *Neuron*, 66(4), 610-618.
- Best, C. T., & Tyler, M. D. (2007). Nonnative and second-language speech perception: Commonalities and complementarities. *Language experience in second language speech learning: In honor of James Emil Flege*, 13-34.
- Binder, J. R., Frost, J. A., Hammeke, T. A., Cox, R. W., Rao, S. M., & Prieto, T. (1997). Human brain language areas identified by functional magnetic resonance imaging. *The Journal of Neuroscience*, 17(1), 353-362.
- Birdsong, D. (2006). Age and second language acquisition and processing: A selective overview. *Language Learning*, 56(s1), 9-49.
- Boersma, P. (2002). Praat, a system for doing phonetics by computer. *Glott international*, 5(9/10), 341-345.
- Brauer, J., & Friederici, A. D. (2007). Functional neural networks of semantic and syntactic processes in the developing brain. *Journal of cognitive neuroscience*, 19(10), 1609-1623.
- Cabeza, R., & Nyberg, L. (2000). Imaging cognition II: An empirical review of 275 PET and fMRI studies. *Journal of cognitive neuroscience*, 12(1), 1-47.
- Carey, S., & Bartlett, E. (1978). Acquiring a single new word.
- D'Anna, C. A., Zechmeister, E. B., & Hall, J. W. (1991). Toward a meaningful definition of vocabulary size. *Journal of Literacy Research*, 23(1), 109-122.

Damasio, H., Grabowski, T. J., Tranel, D., Hichwa, R. D., & Damasio, A. R. (1996). A neural basis for lexical retrieval. *Nature*.

Davis, M. H., & Gaskell, M. G. (2009). A complementary systems account of word learning: neural and behavioural evidence. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, 364(1536), 3773-3800.

Dehaene, S., Dupoux, E., Mehler, J., Cohen, L., Paulesu, E., Perani, D., ... & Le Bihan, D. (1997). Anatomical variability in the cortical representation of first and second language. *Neuroreport*, 8(17), 3809-3815.

Deschamps, I., Baum, S. R., & Gracco, V. L. (2014). On the role of the supramarginal gyrus in phonological processing and verbal working memory: Evidence from rTMS studies. *Neuropsychologia*, 53, 39-46.

Dollaghan, C. (1985). Child Meets WordFast Mapping in Preschool Children. *Journal of Speech, Language, and Hearing Research*, 28(3), 449-454.

Evans, A. C., & Brain Development Cooperative Group. (2006). The NIH MRI study of normal brain development. *Neuroimage*, 30(1), 184-202.

Gaskell, M. G., & Dumay, N. (2003). Lexical competition and the acquisition of novel words. *Cognition*, 89(2), 105-132.

Gathercole, S. E. (2006). Nonword repetition and word learning: The nature of the relationship. *Applied Psycholinguistics*, 27(04), 513-543.

Geiser, E., Zaehle, T., Jancke, L., & Meyer, M. (2008). The neural correlate of speech rhythm as evidenced by metrical speech processing. *Journal of Cognitive Neuroscience*, 20(3), 541-552.

- Gernsbacher, M. A., & Kaschak, M. P. (2003). Neuroimaging studies of language production and comprehension. *Annual Review of Psychology*, 54, 91.
- Gershkoff-Stowe, L., & Hahn, E. R. (2007). Fast mapping skills in the developing lexicon. *Journal of Speech, Language, and Hearing Research*, 50(3), 682-697.
- Goldfield, B. A., & Reznick, J. S. (1990). Early lexical acquisition: Rate, content, and the vocabulary spurt. *Journal of child language*, 17(01), 171-183.
- Golestani, N., & Zatorre, R. J. (2004). Learning new sounds of speech: reallocation of neural substrates. *Neuroimage*, 21(2), 494-506.
- Hickok, G., & Poeppel, D. (2007). The cortical organization of speech processing. *Nature Reviews Neuroscience*, 8(5), 393-402.
- Huang, M. X., Mosher, J. C., & Leahy, R. M. (1999). A sensor-weighted overlapping-sphere head model and exhaustive head model comparison for MEG. *Physics in medicine and biology*, 44(2), 423.
- Hull, R., & Vaid, J. (2007). Bilingual language lateralization: A meta-analytic tale of two hemispheres. *Neuropsychologia*, 45(9), 1987-2008.
- James, T. W., & Gauthier, I. (2006). Repetition-induced changes in BOLD response reflect accumulation of neural activity. *Human brain mapping*, 27(1), 37-46.
- Johnson, J. S., & Newport, E. L. (1989). Critical period effects in second language learning: The influence of maturational state on the acquisition of English as a second language. *Cognitive psychology*, 21(1), 60-99.
- Johnson, M. H. (2001). Functional brain development in humans. *Nature Reviews Neuroscience*, 2(7), 475-483.

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.

Kimppa, L., Kujala, T., & Shtyrov, Y. (2016). Individual language experience modulates rapid formation of cortical memory circuits for novel words. *Scientific Reports*, 6.

KorpusDK. (n.d.). Retrieved May 30, 2016, from http://ordnet.dk/korpusdk_en

Leminen, A., Leminen, M., Lehtonen, M., Nevalainen, P., Ylinen, S., Kimppa, L., ... & Kujala, T. (2011). Spatiotemporal dynamics of the processing of spoken inflected and derived words: a combined EEG and MEG study. *Frontiers in human neuroscience*, 5, 66.

Liikkanen, L. A., Tiitinen, H., Alku, P., Leino, S., Yrttiaho, S., & May, P. J. (2007). The right-hemispheric auditory cortex in humans is sensitive to degraded speech sounds. *Neuroreport*, 18(6), 601-605.

Lin, F. H., Witzel, T., Ahlfors, S. P., Stufflebeam, S. M., Belliveau, J. W., & Hämäläinen, M. S. (2006). Assessing and improving the spatial accuracy in MEG source localization by depth-weighted minimum-norm estimates. *Neuroimage*, 31(1), 160-171.

Lindell, A. K. (2006). In your right mind: Right hemisphere contributions to language processing and production. *Neuropsychology review*, 16(3), 131-148.

López-Barroso, D., Cucurell, D., Rodríguez-Fornells, A., & de Diego-Balaguer, R. (2016). Attentional effects on rule extraction and consolidation from speech. *Cognition*, 152, 61-69.

Newman, S. D., & Twieg, D. (2001). Differences in auditory processing of words and pseudowords: An fMRI study. *Human brain mapping*, 14(1), 39-47.

- Pantazis, D., Joshi, A., Jiang, J., Shattuck, D. W., Bernstein, L. E., Damasio, H., & Leahy, R. M. (2010). Comparison of landmark-based and automatic methods for cortical surface registration. *Neuroimage*, 49(3), 2479-2493.
- Paulesu, E., Vallar, G., Berlingeri, M., Signorini, M., Vitali, P., Burani, C., ... & Fazio, F. (2009). Supercalifragilisticexpialidocious: how the brain learns words never heard before. *Neuroimage*, 45(4), 1368-1377.
- Richards, J. E., & Xie, W. (2015). Chapter One-Brains for All the Ages: Structural Neurodevelopment in Infants and Children from a Life-Span Perspective. *Advances in child development and behavior*, 48, 1-52.
- Richards, J. E., Sanchez, C., Phillips-Meek, M., & Xie, W. (2016). A database of age-appropriate average MRI templates. *Neuroimage*, 124, 1254-1259.
- Rogers, T. T., Lambon Ralph, M. A., Garrard, P., Bozeat, S., McClelland, J. L., Hodges, J. R., & Patterson, K. (2004). Structure and deterioration of semantic memory: a neuropsychological and computational investigation. *Psychological review*, 111(1), 205.
- Posner, M. I., & Rothbart, M. K. (2007). Research on attention networks as a model for the integration of psychological science. *Annu. Rev. Psychol.*, 58, 1-23.
- Ptak, R. (2012). The frontoparietal attention network of the human brain: action, saliency, and a priority map of the environment. *The Neuroscientist*, 18(5), 502-515.
- Sanchez, C. E., Richards, J. E., & Almli, C. R. (2012). Age-specific MRI templates for pediatric neuroimaging. *Developmental neuropsychology*, 37(5), 379-399.
- Shtyrov, Y., Kujala, T., Ilmoniemi, R. J., & Näätänen, R. (1999). Noise affects speech-signal processing differently in the cerebral hemispheres. *NeuroReport*, 10(10), 2189-2192.

Shtyrov, Y., Nikulin, V. V., & Pulvermüller, F. (2010). Rapid cortical plasticity underlying novel word learning. *The Journal of Neuroscience*, 30(50), 16864-16867.

Shtyrov, Y., Kimppa, L., Pulvermüller, F., & Kujala, T. (2011). Event-related potentials reflecting the frequency of unattended spoken words: A neuronal index of connection strength in lexical memory circuits?. *Neuroimage*, 55(2), 658-668.

Shtyrov, Y. (2012). Neural bases of rapid word learning. *The Neuroscientist*, 18(4), 312-319.

Shum, M., Shiller, D. M., Baum, S. R., & Gracco, V. L. (2011). Sensorimotor integration for speech motor learning involves the inferior parietal cortex. *European Journal of Neuroscience*, 34(11), 1817-1822.

Spiegel, C., & Halberda, J. (2011). Rapid fast-mapping abilities in 2-year-olds. *Journal of experimental child psychology*, 109(1), 132-140.

Tadel, F., Baillet, S., Mosher, J. C., Pantazis, D., & Leahy, R. M. (2011). Brainstorm: a user-friendly application for MEG/EEG analysis. *Computational intelligence and neuroscience*, 2011, 8.

Uppenkamp, S., Johnsrude, I. S., Norris, D., Marslen-Wilson, W., & Patterson, R. D. (2006). Locating the initial stages of speech–sound processing in human temporal cortex. *Neuroimage*, 31(3), 1284-1296.

Weber-Fox, C. M., & Neville, H. J. (1996). Maturational constraints on functional specializations for language processing: ERP and behavioral evidence in bilingual speakers. *Journal of cognitive neuroscience*, 8(3), 231-256.

Wechsler, D. (2004). *The Wechsler intelligence scale for children—fourth edition*. Pearson Assessment, London.

Weiss-Croft, L. J., & Baldeweg, T. (2015). Maturation of language networks in children: A systematic review of 22 years of functional MRI. *NeuroImage*, 123, 269-281.

Vigneau, M., Beaucoisin, V., Hervé, P. Y., Jobard, G., Petit, L., Crivello, F., ... & Tzourio-Mazoyer, N. (2011). What is right-hemisphere contribution to phonological, lexico-semantic, and sentence processing?: Insights from a meta-analysis. *Neuroimage*, 54(1), 577-593.

Wilson, S. M., & Iacoboni, M. (2006). Neural responses to non-native phonemes varying in producibility: Evidence for the sensorimotor nature of speech perception. *Neuroimage*, 33(1), 316-325.

Yue, J., Bastiaanse, R., & Alter, K. (2014). Cortical plasticity induced by rapid Hebbian learning of novel tonal word forms: Evidence from mismatch negativity. *Brain and language*, 139, 10-22.

Zhao, J., Shu, H., Zhang, L., Wang, X., Gong, Q., & Li, P. (2008). Cortical competition during language discrimination. *NeuroImage*, 43(3), 624-633

Supplementary data

Supplementary data 1. Relative response strengths at 91 ms (first peak after disambiguation point)

for novel word forms with non-native phonology (top panel, yellow) and musical rain sounds (bottom panel, gray). Asterisks denote statistical significance: *: $p < 0.05$; #: $p < 0.1$.

Supplementary data 2. Relative response strengths at 317 ms (first peak after disambiguation point) for real native language word forms (top panel, cyan) and novel native language word forms (bottom panel, magenta). Asterisks denote statistical significance: *: $p < 0.05$; #: $p < 0.1$.

	Age (months)	Reading classification	PRI	WMI	VCI	Digit span forward	Digit span backward
NLWF, left	$\rho=-0.085$, $p=0.700$	$\rho=-0.081$, $p=0.715$	$\rho=0.033$, $p=0.881$	$\rho=-0.006$, $p=0.978$	$\rho=-0.076$, $p=0.730$	$\rho=-0.115$, $p=0.603$	$\rho=-0.185$, $p=0.398$
NLWF, right	$\rho=-0.358$, $p=0.094$	$\rho=-0.259$, $p=0.233$	$\rho=-0.154$, $p=0.483$	$\rho=0.066$, $p=0.765$	$\rho=0.116$, $p=0.598$	$\rho=-0.380$, $p=0.074$	$\rho=-0.156$, $p=0.476$
NNWF, left	$\rho=0.059$, $p=0.788$	$\rho=0.195$, $p=0.371$	$\rho=0.283$, $p=0.190$	$\rho=0.060$, $p=0.786$	$\rho=-0.094$, $p=0.671$	$\rho=0.274$, $p=0.206$	$\rho=-0.258$, $p=0.235$
NNWF, right	$\rho=-0.064$, $p=0.771$	$\rho=0.069$, $p=0.756$	$\rho=0.102$, $p=0.642$	$\rho=0.053$, $p=0.809$	$\rho=-0.072$, $p=0.744$	$\rho=-0.226$, $p=0.299$	$\rho=-0.001$, $p=0.996$
NSS, left	$\rho=-0.174$, $p=0.427$	$\rho=-0.199$, $p=0.363$	$\rho=-0.082$, $p=0.710$	$\rho=0.103$, $p=0.604$	$\rho=0.098$, $p=0.655$	$\rho=0.207$, $p=0.344$	$\rho=-0.263$, $p=0.226$
NSS, right	$\rho=0.015$, $p=0.946$	$\rho=0.063$, $p=0.774$	$\rho=-0.061$, $p=0.782$	$\rho=0.089$, $p=0.687$	$\rho=-0.033$, $p=0.880$	$\rho=-0.043$, $p=0.844$	$\rho=0.080$, $p=0.716$

Supplementary data 3. Correlation of performance measures and ERF variables. All correlation analyses are conducted using Spearman's correlation, N=23 for all analyses. The scale for reading classification is as follows: 0 if the child could not read, 1 if the child made more than 5 mistakes when reading a 45-word story, and 2 if child made less than 5 errors in reading. PRI is assessed using block design subtest, WMI with digit span subtest, and VCI with vocabulary subtest of Danish WISC-IV; standard scores are used in correlation analyses. Digit span is measured as largest number of digits the child could repeat either forward or backward. ERF measures (leftmost column) are relative changes in ERF amplitudes from initial to final part of the experiment, separately for left and right

hemispheres. Following abbreviations are used: NLWF, unfamiliar native language word form; NNWF, non-native word form; NSS, non speech sound.