Unlocking adults' implicit statistical learning by cognitive depletion

Eleonore H. M. Smalle[a,1], **Tatsuya Daikoku**[b], **Arnaud Szmalec**[a,c,d], **Wouter Duyck**[a], **and Riikka Möttönen**[e,f]

[a]Department of Experimental Psychology, Ghent University, 9000 Ghent, Belgium; [b]International Research Center for Neurointelligence, The University of Tokyo, Tokyo 158-8557, Japan; [c]Psychological Research Institute, Université catholique de Louvain, 1348 Ottignies-Louvain-la-Neuve, Belgium; [d]Institute of Neuroscience, Université catholique de Louvain, 1348 Ottignies-Louvain-la-Neuve, Belgium; [e]School of Psychology, University of Nottingham, Nottingham NG7 2QL, United Kingdom; and [f]Cognitive Science, Department of Digital Humanities, University of Helsinki, 00014 Helsinki, Finland

Edited by Richard Aslin, Haskins Laboratories, New Haven, CT; received December 17, 2020; accepted November 4, 2021

Human learning is supported by multiple neural mechanisms that maturate at different rates and interact in mostly cooperative but also sometimes competitive ways. We tested the hypothesis that mature cognitive mechanisms constrain implicit statistical learning mechanisms that contribute to early language acquisition. Specifically, we tested the prediction that depleting cognitive control mechanisms in adults enhances their implicit, auditory word-segmentation abilities. Young adults were exposed to continuous streams of syllables that repeated into hidden novel words while watching a silent film. Afterward, learning was measured in a forced-choice test that contrasted hidden words with nonwords. The participants also had to indicate whether they explicitly recalled the word or not in order to dissociate explicit versus implicit knowledge. We additionally measured electroencephalography during exposure to measure neural entrainment to the repeating words. Engagement of the cognitive mechanisms was manipulated by using two methods. In experiment 1 ($n = 36$), inhibitory theta-burst stimulation (TBS) was applied to the left dorsolateral prefrontal cortex or to a control region. In experiment 2 ($n = 60$), participants performed a dual working-memory task that induced high or low levels of cognitive fatigue. In both experiments, cognitive depletion enhanced word recognition, especially when participants reported low confidence in remembering the words (i.e., when their knowledge was implicit). TBS additionally modulated neural entrainment to the words and syllables. These findings suggest that cognitive depletion improves the acquisition of linguistic knowledge in adults by unlocking implicit statistical learning mechanisms and support the hypothesis that adult language learning is antagonized by higher cognitive mechanisms.

auditory statistical learning | implicit learning | electroencephalography | cognitive load | transcranial magnetic stimulation

Human learning is thought to be supported by the interactions between two basic memory systems of the brain, namely declarative and nondeclarative memory (1). Declarative memory is characterized by voluntary, explicit, attention-based processes, such as recall and recognition of facts/events, and is mediated by medial-temporal lobe and prefrontal cortex structures (2). Nondeclarative memory, also referred to as procedural memory, on the other hand is part of implicit memory and includes the acquisition of a heterogeneity of skills, habits, and procedures. It is mediated by basal ganglia, cerebellar, and neocortical structures, as well as parts of the prefrontal cortex [e.g., Broca's area (3–5)].

Accumulating evidence supports a competitive relationship between these two memory systems during human skill learning. Suppression of the declarative memory system by interventions like repetitive transcranial magnetic stimulation (TMS), distraction tasks, alcohol consumption, hypnosis, intake of benzodiazepines, or cognitive fatigue, can actually enhance performance in implicit, perceptual-motor learning tasks such as the serial-reaction time task (6–11) or intuitive reasoning tasks (12). These findings suggest that higher-level cognitive functions associated with declarative memory and supported by the prefrontal cortex can interfere with behavior that is naturally driven by implicit learning processes (13). However, it remains unresolved whether competing memory systems also affect implicit statistical learning abilities that are critical for the early, rapid acquisition of language in infants (14). This is an important question, as it could explain why infants and children pick up languages with less effort than adults (cf "What don't we know?") (15).

Language acquisition involves many different memory and learning processes that are dependent on both procedural and declarative memory (2, 16). The first step for infants acquiring language is to gain knowledge about the phonological structure in one's spoken language system, the probabilistic constraints on how speech sounds combine (i.e., phonotactic learning), and the segments of continuous speech (i.e., word forms) (17). Word form learning takes place already in the first 12 months of life and is an important precursor to vocabulary acquisition (i.e., mapping form to meaning) and more complex language acquisition (e.g., grammar) later in development (18). In the present study, we focus on statistical learning mechanisms that contribute to word segmentation and thus novel word form learning in the early stages of language acquisition.

Statistical learning is generally known as the ability to pick up on patterns in the environment through extraction of

Significance

Statistical learning mechanisms enable extraction of patterns in the environment from infancy to adulthood. For example, they enable segmentation of continuous speech streams into novel words. Adults typically become aware of the hidden words even when passively listening to speech streams. It remains poorly understood how cognitive development and brain maturation affect implicit statistical learning (i.e., infant-like learning without awareness). Here, we show that the depletion of the cognitive control system by noninvasive brain stimulation or by demanding cognitive tasks boosts adults' implicit but not explicit word-segmentation abilities. These findings suggest that the adult cognitive architecture constrains statistical learning mechanisms that are likely to contribute to early language acquisition and opens avenues to enhance language-learning abilities in adults.

Author contributions: E.H.M.S. and R.M. designed research; E.H.M.S. performed research; E.H.M.S., T.D., and R.M. analyzed data; W.D. helped in funding of the first author; R.M. helped in funding of the research (tools and participants); E.H.M.S. and R.M. wrote the paper; and A.S. and W.D. reviewed and edited the paper.

The authors declare no competing interest.

This article is a PNAS Direct Submission.

[1] To whom correspondence may be addressed. Email: eleonore.smalle@ugent.be.

This article contains supporting information online at http://www.pnas.org/lookup/suppl/doi:10.1073/pnas.2026011119/-/DCSupplemental.

frequent regularities and distributional properties. The term was first introduced in the field of cognitive psychology by the work of Saffran, Aslin, and Newport (1996) (19), who demonstrated that infants of only 8-mo-old can extract word boundaries and segment novel word forms from a continuous stream of speech sounds with no other cue than the transitional probabilities between syllables. Later, this learning was also demonstrated in older children and adults (20, 21) and across different domains (e.g., music and grammar) or modalities (e.g., auditory, visual, and motor) (22, 23), indicating that statistical learning is a largely continuous and domain-general learning mechanism for skill acquisition across the human life span.

In a typical statistical learning experiment, participants are repeatedly exposed to patterned stimuli such as consonant strings from an artificial grammar, or recurrent syllable triplets. Learning is then typically assessed postexposure by using a two-alternative forced-choice recognition task in which triplets from the exposure stream are pitted against foils. Participants have to indicate which of the two triplets sounded most familiar, and above-chance accuracy is taken as indication of learning. Since statistical learning occurs without any instruction or intention to learn, it is often assumed to result in implicit memory representations (24). This view is also supported by the evidence that statistical learning occurs in infants and even in sleeping neonates (25). However, in recent work, Batterink and colleagues demonstrated that even without intention to learn, adults acquire mainly explicit knowledge of the novel word forms during statistical learning (26–29). This can be derived from the observation that participants' performance was above chance when they were confident remembering the triplet but at chance when they were unconfident. Knowledge is implicit when participants lack awareness of what they have learned. This means that if participants perform also above chance when they are unconfident, knowledge is inferred to be implicit (30). In contrast, if they perform at chance level when confidence is low, no implicit knowledge is gained. Although statistical learning may produce additional implicit knowledge that cannot be assessed by the recognition and memory judgement tasks (e.g., ref. 28), Batterink's earlier findings show that adults store the acquired word knowledge mainly in the explicit memory system.

We and others have proposed that cognitive development and maturation of the prefrontal areas negatively affect language acquisition, such as word form or grammar learning (31–35). For instance, we showed that children outperform adults on the Hebb repetition learning paradigm (32, 33), a memory paradigm in which participants are asked to immediately recall syllable sequences that consist of hidden repeated word forms. Interestingly, in a follow-up study, we found that cognitive depletion by TMS to the left dorsolateral prefrontal cortex (DLPFC), an area closely related to declarative memory and cognitive control, enhanced Hebb performance in adult participants (34). This suggests that late-developing prefrontal cognitive mechanisms can induce changes in efficiently acquiring sequential language information from the environment, a finding that is largely in line with previously reported evidence in skill learning (13). Recently, we corroborated this idea further by showing enhanced phonotactic constraint learning in adults under cognitive fatigue (35). Based on these findings, we hypothesize that the higher cognitive control system could reduce access to implicit memory processes in adults, thereby making them less efficient in language acquisition relative to infants and children. This idea is in line with the well-known less-is-more hypothesis that attributes developmental changes in language acquisition, such as phonology and grammar, to maturational changes in attention and memory capacities (36–38). In our previous work, participants were explicitly asked to memorize (34) or produce (35) syllable sequences and thus exposure to the novel language was not passive, or "infant like." Moreover, we did not separate implicit and explicit memory representations. Thus it remains unresolved how higher-order cognitive functions affect acquisition of implicit linguistic knowledge during passive listening to continuous speech using statistical learning mechanisms that support infant language acquisition (23, 39).

The aim of the current study was to directly address this question using the auditory statistical learning paradigm. In particular, we aimed to determine whether a temporary depletion of the higher cognitive control system, using two different interventions, can unlock adults' implicit statistical learning processes that serve infant word segmentation. To investigate this, we exposed young adults to continuous streams of syllables with, unknown to them, repeating three-syllable pseudowords, while watching a silent film. In the first experiment, inhibitory continuous theta-burst stimulation was used to induce a long-lasting disruption in left DLPFC or a control site prior to exposure, similar to the method used in Smalle et al., 2017 (34). In the second experiment, participants first performed an effortful dual working-memory task under high– or low–cognitive-load(HCL and LCL, respectively) conditions, which induces cognitive fatigue that hampers subsequent cognitive performance (7, 35, 40), or did not perform a cognitive load task prior to the language exposure (control or no-load condition). Our primary measure of statistical learning was the offline recognition of the hidden words, which was assessed 15 min after exposure. This was combined with a memory judgement procedure, which measured how confident the participants were that they remembered the hidden words. This task dissociates explicit versus implicit memory representations (e.g., refs. 27–29, 41). In both experiments, electroencephalography (EEG) was also measured during the 20-min language exposure in order to investigate an online perceptual component as second independent measure of statistical learning. Research has shown that the steady-state response of the brain shows a decrease at the frequency of individual syllables and an increase at the rhythm of three-syllable words while listening to continuous sound streams that consist of repeating three-syllable structures. This shift in neural entrainment indicates online statistical learning of novel words as a function of auditory exposure (29). Overall, we predicted that TMS-induced disruption of the DLPFC (in experiment 1) and cognitive fatigue (in experiment 2) would enhance statistical language learning and especially strengthen implicit memory representations for the hidden novel words.

Results

The Effect of Cognitive Depletion on Recognition of the Hidden Words.

Experiment 1. Participants performed above chance on the forced-choice recognition task, indicating statistical learning, in both groups (i.e., DLPFC: mean = 68.8, SE = 3.5, $t_{17} = 5.4$, $P < 0.001$, $d = 1.3$; Vertex: mean = 57.3, SE = 3.4, $t_{17} = 2.2$, $P < 0.05$, $d = 0.5$). TMS-induced disruption of the DLPFC improved recognition accuracy of the hidden words [the effect of TMS: β = 0.28, SE = 0.096, $Z = 2.87$; X^2 (1) = 8.25, $P = 0.004$, $d = 0.6$, Fig. 1]. For the unconfident responses (64% of all trials; 31% in TMS-disrupted group, 33% in control group), accuracy was significantly above chance in the disrupted group (mean = 63.9, SE = 3.7, $t_{17} = 3.8$, $P < 0.001$, $d = 0.9$) but not in the control group (mean = 51.7, SE = 3.8, t < 1, $P = 0.33$, $d = 0.1$). The control and disrupted groups differed significantly from each other [the main effect of TMS: β = 0.23, SE = 0.108, $Z = 2.16$, X^2 (1) = 4.68, $P = 0.030$, $d = 0.5$]. For the confident responses (36% of all trials; 19% in TMS-disrupted group, 17% in control group), all participants showed above-chance accuracy (DLPFC: mean = 75.9, SE = 6.0, $t_{17} = 4.3$, $P < 0.001$,

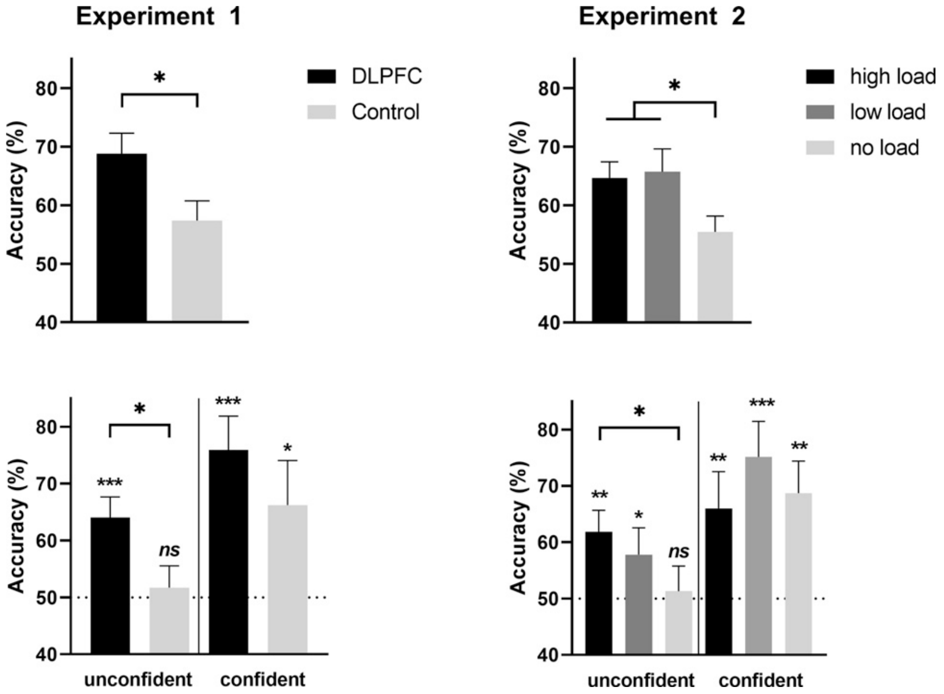

Fig. 1. Behavioral results reflecting statistical word segmentation. The *Upper* panel shows total recognition accuracy (%) for the hidden words as a function of the cognitive disruption in experiment 1 (*Left*) and experiment 2 (*Right*). The *Lower* panel shows recognition accuracy for the hidden words when participants indicated low and high confidence in remembering the hidden word, referred to as unconfident versus confident responses, respectively. Above-chance performance on confident responses reflects knowledge stored in explicit memory. Above-chance performance on unconfident responses reflects knowledge stored in implicit memory. Error bars denote SEMs. Asterisks denote significance for one-tailed *t* tests: *$P < 0.05$, **$P < 0.01$, **$P < 0.001$. *ns*, nonsignificant.

$d = 1.0$; Vertex: mean = 66.3, SE = 7.8, $t_{16} = 2.08$, $P < 0.05$, $d = 0.5$). There was no significant difference between groups [the main effect of TMS: $\beta = 0.35$, SE = 0.214, $Z = 1.62$, X^2 (1) = 2.63, $P = 0.11$].

Experiment 2. Participants performed above chance on the forced-choice recognition task, indicating statistical learning, in all groups (high load: mean = 64.7, SE = 2.7, $t_{19} = 5.4$, $P < 0.001$, $d = 1.2$; low load: mean = 65.6, SE = 3.9, $t_{19} = 4.0$, $P < 0.001$, $d = 0.9$; no load: mean = 55.3, SE = 2.7, $t_{19} = 1.9$, $P < 0.05$, $d = 0.4$). Cognitive load improved recognition accuracy of the hidden words [effect of cognitive load: high versus Control: $\beta = 0.42$, SE = 0.19, $Z = 2.23$, $P = 0.026$, $d = 0.4$; low versus Control: $\beta = 0.46$, SE = 0.19, $Z = 2.47$, $P = 0.013$, $d = 0.5$; X^2 (2) = 7.50, $P = 0.024$, Fig. 1]. For the unconfident responses (64% of all trials; 21% in high, 20% in low, and 24% in no-load group), accuracy was above chance in the high–cognitive-load group (mean = 61.8, SE = 3.8, $t_{19} = 3.08$, $P < 0.01$, $d = 0.7$) and in the low–cognitive-load group (mean = 57.8, SE = 4.7, $t_{19} = 1.64$, $P = 0.05$, $d = 0.4$) but not in the no cognitive load group (mean = 51.3, SE = 4.5, $t < 1$, $P = 0.4$, $d = 0.07$). The cognitive load enhanced accuracy relative to the control group [high versus Control: $\beta = 0.52$, SE = 0.20, $Z = 2.57$, $P = 0.01$, $d = 0.5$; low versus Control: $\beta = 0.37$, SE = 0.20, $Z = 1.81$, $P = 0.070$, $d = 0.4$; the main effect of cognitive load: X^2 (2) = 7.16, $P = 0.028$]. For the confident responses (36% of all trials; 12% in high, 14% in low, and 10% in no-load group), all groups showed above-chance accuracy (high cognitive load: mean = 66.0, SE = 6.5, $t_{19} = 2.5$, $P = 0.012$, $d = 0.6$; low cognitive load: mean = 75.1, SE = 6.3, $t_{19} = 4.0$, $P < 0.001$, $d = 0.9$; no cognitive load: mean = 68.7, SE = 5.7, $t_{19} = 3.3$, $P < 0.01$, $d = 0.7$). No significant differences were found between the groups [high versus Control: $\beta = -0.012$, SE = 0.45, $Z = -0.028$, $P = 0.98$; low versus Control: $\beta = 0.43$, SE = 0.46, $Z = 0.94$, $P = 0.35$; the main effect of cognitive load: X^2 (2) = 1.24, $P = 0.54$].

The Effect of Cognitive Depletion on Neural Entrainment during Exposure.

Experiment 1. During exposure, neural entrainment to the underlying word patterns was enhanced during structured exposure relative to random exposure: intertrial coherence (ITC) increased at the word frequency ($P = 0.01$, Cohen's $d = 1.1$) but decreased at the syllable frequency ($P = 0.003$, $d = 1.4$) relative to exposure to the random syllable stream [Frequency × Exposure: F (1, 31) = 16.4, $P < 0.001$, Fig. 2]. Disrupting the left DLPFC with TMS increased the neural word-learning index (WLI) [Exposure: F (1, 30) = 31.6, $P < 0.001$, TMS × Exposure: F (1, 30) = 5.6, $P = 0.025$, random: $P = 0.6$, $d = 0.2$, structured: $P = 0.004$, $d = 1.1$, Fig. 3].

Experiment 2. During exposure, neural entrainment to the underlying word patterns was enhanced during structured exposure relative to random exposure: ITC increased at word frequency ($P < 0.001$, $d = 1.4$) but decreased at syllable frequency ($P = 0.08$, Cohen's $d = 0.5$) relative to exposure to a random syllable sequence [Frequency × Exposure: F (1, 57) = 19.3, $P < 0.001$, Fig. 2]. Cognitive load did not affect the neural WLI [Exposure: F (1, 55) = 40.4, $P < 0.001$, Load × Exposure: F (2, 55) = 1.04, $P = 0.36$, Fig. 3].

Discussion

Overall, our findings provide evidence for a competitive interaction between higher cognitive control functions and implicit statistical learning mechanisms that contribute to word segmentation in the early stages of language acquisition. More specifically, we depleted the cognitive control mechanisms in young adults by applying TMS to the left DLPFC (experiment 1) and by inducing cognitive fatigue (experiment 2). These interventions with long-lasting effects were applied prior to exposure to a continuous stream of speech sounds that consisted of trisyllabic word patterns. There were no acoustic cues of the word boundaries in

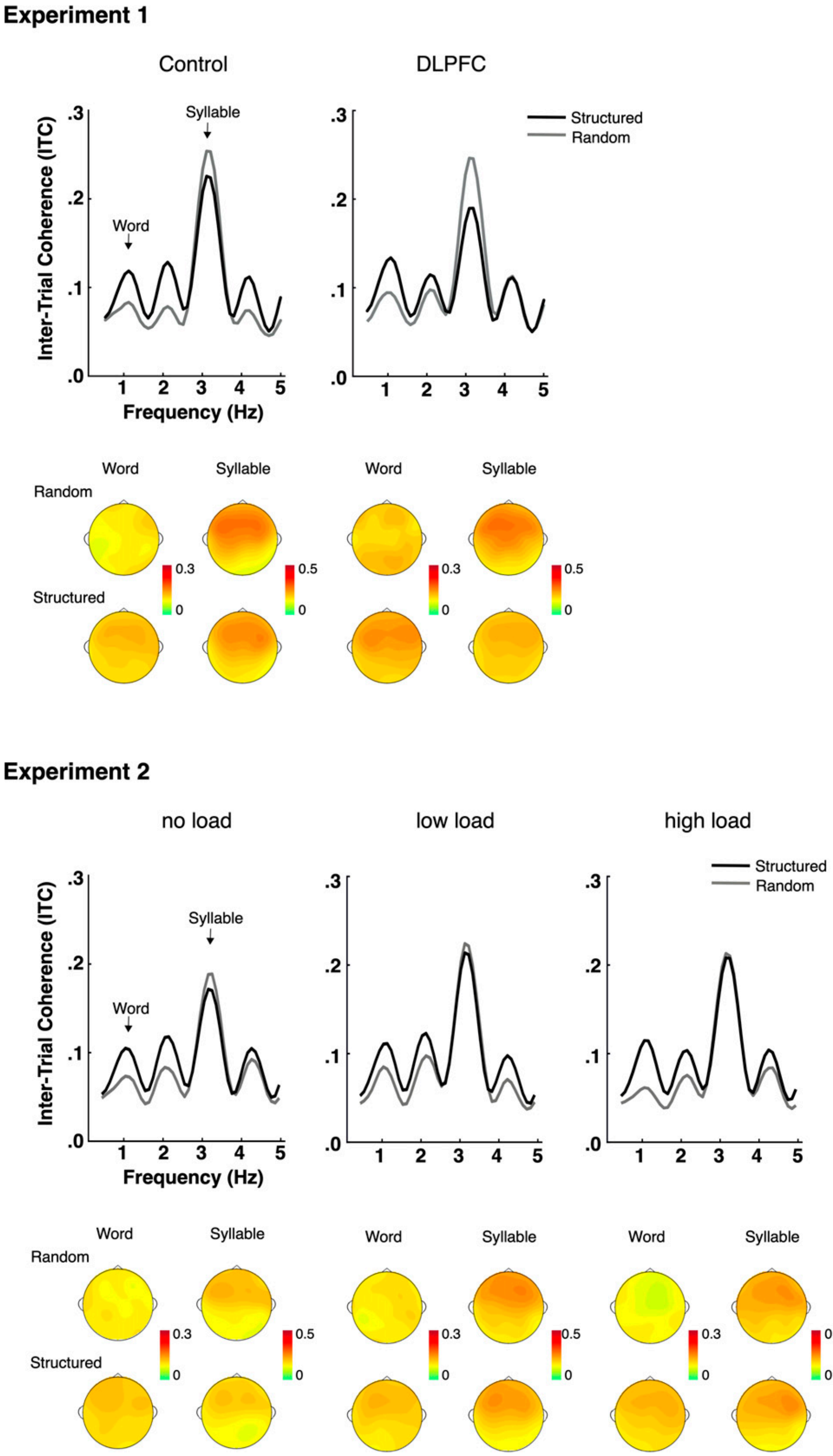

Fig. 2. EEG results reflecting online statistical learning: ITC (mean of six centrofrontal midline electrodes, FC1, FC2, F3, F4, FC5, and Fz) during exposure to random and structured sequences in experiment 1 (TMS) and experiment 2 (Cognitive Load). Topographical plot shows distribution of ITC across the scalp, as a function of exposure and frequency. ITC values were used to calculate the WLI (i.e., ITC word / ITC syllable) (Fig. 3).

the streams. Both TMS-induced disruption of the left DLPFC and cognitive fatigue enhanced recognition accuracy for the hidden words, indicating enhanced statistical learning. Intriguingly, these cognitive manipulations specifically enhanced recognition accuracy when the participants had low confidence in remembering the hidden words, indicating enhanced implicit statistical

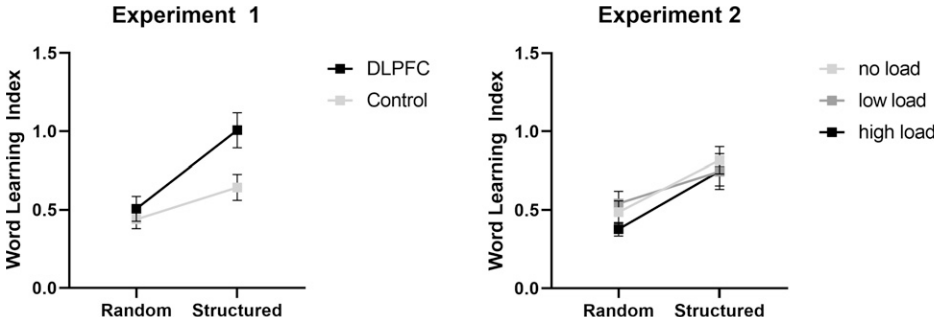

Fig. 3. EEG results reflecting online statistical learning: Changes in WLI as a function of exposure (Random versus Structured sequences) with and without cognitive disruption (experiment 1: TMS; experiment 2: Cognitive load). Error bars denote SEMs.

learning. In fact, the control adults showed chance-level accuracy for the words stored in implicit memory, in agreement with earlier studies in adults (27–29), whereas the cognitively depleted adults showed above-chance–level recognition accuracy. The results therefore suggest that cognitive depletion unlocks implicit word-segmentation abilities in adults.

All groups showed above-chance recognition accuracy when confidence in remembering the hidden words was high, indicating explicit statistical learning. Cognitive manipulations had no effect on explicit statistical learning. This is important because the cognitive manipulations could have reduced explicit or declarative learning mechanisms. Our findings suggest that the enhancement of implicit statistical learning did not occur at the expense of explicit memory formation and thus does not completely support the idea of a direct competition between implicit and explicit learning. These observations are more in line with the idea that implicit statistical learning remains available across development (39) but that the adult brain imposes a “bottleneck,” which prioritizes access to the explicit memory system (42). Disrupting the higher cognitive control mechanisms abolishes this “bottleneck” and, as a consequence, improves adults’ capacity to simultaneously store linguistic knowledge in both implicit and explicit memory systems. This results in overall enhanced word recognition in the disrupted adults. Further research is needed to test this model and investigate how facilitating the higher-order cognitive system affects acquisition of implicit versus explicit linguistic knowledge.

Our experimental paradigm included a break after the exposure to allow participants to recover from the cognitive manipulations before performing the memory tests in experiments 1 and 2. It is possible that this delay strengthened implicit memory consolidation rather than episodic recall so that participants were less likely to explicitly remember the words from the stream. In fact, in both experiments, participants were confident in only 36% of the total trials. The memory judgement (i.e., the proportion of confident versus unconfident responses) was however unaffected by TMS or cognitive fatigue. Cognitive manipulations specifically affected accuracy of the unconfident but not the confident responses after a 15-min delay. Further research is needed to investigate whether this effect is dependent on a short consolidation period and whether a longer consolidation period would further strengthen this effect.

We additionally measured neural entrainment during statistical learning in order to investigate perceptual binding of neighboring syllables into words during passive listening to the structured syllable streams, as a second, independent marker of learning. As expected based on recent findings by Batterink and Paller (2017), neural entrainment decreased at the frequency of syllables, while it increased at the level of the words as a function of the structured exposure (29). This indicates successful perceptual binding across all participants. Interestingly, however, TMS-induced disruption of the left DLPFC enhanced this perceptual binding (measured with a WLI), whereas cognitive fatigue had no effect on it, although both TMS and cognitive fatigue enhanced memory for the words as measured in the postexposure recognition task. These findings are in line with the view that a perceptual binding (also called, “processing-based”) component of statistical learning (as measured here with online EEG) is dissociable from a memory storage or retrieval component of statistical learning (as measured here with the offline recognition task) (26, 43, 44). This is further supported by the absence of correlation between the online and offline measures (*SI Appendix*). Exploratory oscillatory power analyses (*SI Appendix*) showed that TMS-induced disruption of the DLPFC marginally enhanced the overall power of theta (4 to 8 Hz) and alpha oscillations (8 to 12 Hz) and significantly enhanced alpha oscillations in the parietal-occipital area specifically. Cognitive fatigue, in contrast, decreased the overall power of delta (1 to 4 Hz) oscillations. This suggests that the cognitive manipulations had different neural consequences. Theta and lower alpha power are strong in early childhood, and it decreases during cognitive development (e.g., refs. 45, 46). Power of lower-frequency oscillations, such as delta and theta, continue to decrease during adulthood (47). Thus, the pattern of results invites a tentative interpretation that disruption of the left DLPFC made the young adults to function in the same way as young children do and therefore enhanced both perceptual binding and implicit memory components of statistical learning. The cognitive fatigue manipulation, however, possibly mimicked cognitive decline throughout adulthood. Statistical language learning and implicit motor learning have been shown to be well-preserved in older adults (20, 48–51).

The current findings are largely in agreement with previous studies showing that TMS-induced disruption of the DLPFC and/or cognitive fatigue enhances sequence learning in the motor system (6, 7) and the language system (34, 35). The present study demonstrates the effects of cognitive depletion on implicit versus explicit outcomes of statistical language learning in the absence of explicit instructions (i.e., passive listening during exposure) and replicating these effects by using two different methods: TMS-induced disruption and cognitive fatigue. A previous study found that only high cognitive load induced feelings of cognitive fatigue and improved motor sequence learning (7). We found that both low- and high-load tasks increased feelings of cognitive fatigue in our participant sample, and consequently, statistical language learning was enhanced in both groups (relative to a control condition where no cognitive load task was performed).

Statistical language learning relies on multiple brain regions and pathways that work in parallel, such as hippocampus (44), inferior frontal cortex (52), striatum (53), and auditory–motor pathways (54). It remains, however, poorly understood how these brain regions and pathways contribute to acquisition of implicit versus explicit memory representations for linguistic knowledge from structured sound sequences. Another important goal for future research is to determine how the prefrontal cognitive control mechanisms affect competition and cooperation between various language-learning mechanisms.

This experimental study provides causal evidence for a hypothesis that the cognitive control system constrains implicit language-learning abilities in adults. Our findings show that depletion of the mature cognitive system can enhance implicit, statistical learning mechanisms that are used in early language acquisition. This finding is an important step in science of human language development, as it could help us to understand maturational constraints and interindividual differences (e.g., language-related difficulties) in language learning. Importantly, cognitive depletion could be a key for unlocking infant-like implicit learning mechanisms and, as a result, enhance foreign language learning in adults.

Materials and Methods

Participants. We decided to test 20 participants in each group based on the large effect sizes in a previous study by Batterink and colleagues (28), who used similar dependent measures, and in our previous study, in which we used an identical TMS paradigm as in the current study (34). In our previous study, we obtained a t test effect size d of 0.88 for the difference in Hebb learning performance between a DLPFC-disrupted group (n = 14, mean correct recall on the last block of Hebb trials = 89.3%, SD = 16.04%) and a control group (n = 14, mean correct recall = 72.5%, SD = 21.9%). After postcollection exclusion of four nonfluent English–speaking participants, we report the data of 36 participants in experiment 1 who were randomly assigned to either the left DLPFC stimulation (n = 18, age = $25.3_M \pm 4.8_{SD}$, 9 females) or the control stimulation to Vertex (n = 18, age = $23.4_M \pm 5.0_{SD}$, 12 females). In three participants (two in the TMS group and one in the control group), EEG was not recorded due to technical failures. Data of these participants were however still included for the behavioral analyses. In experiment 2, we report the data of 60 participants who were randomly assigned to a high-load (n = 20, age = $22.7_M \pm 3.1_{SD}$, 12 females), low-load (n = 20, age = $21.8_M \pm 3.9_{SD}$, 16 females), or no-load (i.e., control, n = 20, age = $18.8_M \pm 0.81_{SD}$, 16 females) condition. We have no EEG recording from two participants (one in the control group and one in the low load group) due to technical failure, but their behavioral data were included in the analyses. All included participants were right-handed and native (or nonnative but fluent) English speakers. None of the participants had a history of language (learning) impairments or neurological problems. Participants from all groups were matched on various cognitive control abilities (Table 1). Experiments were undertaken with written informed consent and blind to the purpose of the study. Participants received financial compensation at the end of the experiment (£10/h). The study was approved by the Research Ethics Committee of the School of Psychology at the University of Nottingham (reference: F1003).

Experimental Design. During a pretest, individual cognitive control abilities were assessed (Table 1). Participants in experiment 2 were additionally pretested on their maximal processing speed capacity (i.e., the shortest processing time interval to simultaneously perform two working memory tasks while maintaining an accuracy of at least 85%). This assessment was necessary for the experimental cognitive load manipulation in experiment 2 [*Cognitive Fatigue (Experiment 2)*]. For both experiments, the main experiment took place on a separate day, during which the participants in both experiments were exposed to a 20-min auditory syllable stream while EEG was recorded. After a 15-min break, the participants completed a postexposure recognition test that examined implicit and explicit memory of the hidden words. The main experimental design is presented in Fig. 4.

TMS (Experiment 1). TMS was delivered using a 70-mm–diameter figure-eight coil attached to DuoMAG XT stimulator (by Deymed, Brainbox Ltd). We first localized the left DLPFC in each participant using the BeamF3 algorithm (55, 56). We then identified the left motor cortex as the spot eliciting reliable twitches in the resting contralateral hand. The active motor threshold (aMT) was defined as the lowest intensity at which TMS elicited at least 5 out of 10 visible muscle twitches, while the subject sustained a light contraction of their pinch. After defining the participant's aMT, the coil was placed over the left DLPFC or a control area (2 cm posterior to vertex), similar to our previous TMS study (34). The control area was assumed not to play a role in statistical learning or cognitive control (57). The location of the coil (i.e., DLPFC or Vertex) was decided randomly based on the number of participants entering the experiment. The coil was placed tangentially to the scalp with the handle pointing posterior at a 45° angle with respect to the anterior–posterior axis for DLPFC and at 0° for the control site. The intensity of the stimulation was set at 80% of each participant's aMT [i.e., at 43.7% (SD = 6.3) for the DLPFC group and at 49.3% (SD = 8.3) for the control group]. Similar to Smalle et al., 2017 (34), a modified continuous theta-burst stimulation (cTBS) protocol was used in which 600 pulses were delivered in a continuous train of 200 bursts. Each burst consisted of three pulses at 30 Hz, repeated at 6 Hz. The total stimulation duration was 30 s. This modified cTBS protocol is known to inhibit cortical excitability for at least 30 min after stimulation over the primary motor area (58). Importantly, cTBS to the DLPFC does not impair metacognition or conscious perception processes (59).

Cognitive Fatigue (Experiment 2). Cognitive fatigue was induced with the TloadDback task (40). The script of the TloadDback task is freely available on Open Science Framework (https://osf.io/ay6er). The task was run in Matlab2016b/Psychtoolbox on a Dell laptop (refresh rate, 60 Hz). The letters were centrally presented in Arial, font size 120, on a 15.6-inch screen. For each participant, the shortest time needed for accurately processing two ongoing task demands, namely n-back letter detection and parity number decision, was defined during a pretest on a first assessment day (*Experimental Design*). During the TloadDback task, digits (1 to 4 and 6 to 9) and letters (A, C, T, L, N, E, U, and P) were presented in alternation on the screen. Participants were instructed to press the space bar with their left hand every time the displayed letter was the same as the last-seen letter and to indicate with their right hand whether the subsequently displayed digit was odd (pressing "1" on the numeric keypad) or even (pressing "2"). Different levels of cognitive load were created by presenting the two tasks at different paces based on our participant's pretested maximum processing speed capacities (no a priori group differences; Table 1). This is defined as the fastest stimulus time duration (STD) allowing an accuracy performance of at least 85%. Under HCL conditions, the task was performed for 16 min at the subject's max. STD while under LCL conditions, the presentation rate was made one-third slower (i.e., STD = max. STD + 1/2 max. STD). This results in different cognitive demands, with higher sustained attentional-control requirements for the former condition (despite the same level of task complexity), eventually leading to a higher state of "cognitive depletion or mental fatigue" (7, 40). As expected, the LCL participants showed higher dual-task performance than the HCL participants, who performed around the minimal 85% accuracy level defined during the pretest

Table 1. Participant characteristics: Pretested individual cognitive control abilities of the different groups

	Control	Depleted	
Experiment 1	Vertex n = 18	DLPFC N = 18	
Digit span (forward + backward)	18 (min = 11, max = 25)	19 (min = 11, max = 28)	
WCST (N perseveration errors)	7 (min = 5, max = 17)	7 (min = 5, min = 11)	
Experiment 2	No load (N = 20)	LCL (N = 20)	HCL (N = 20)
Digit span (forward + backward)	18 (10 to 22)	19 (12 to 24)	20 (12 to 29)
ECST (N perseveration errors)	8 (5 to 19)	8 (5 to 17)	7 (5 to 11)
STD (in seconds)	0.91 (0.46 to 1.3)	0.87 (0.50 to 1.4)	0.91 (0.46 to 1.3)

Values represent average scores with minimum to maximum in parentheses. WCST = Winsconsin Card Sorting Test for cognitive reasoning; STD = (pre-tested) Stimulus Time Duration or maximal processing capacity (i.e., the shortest processing time interval to simultaneously perform two working memory tasks while maintaining an accuracy of at least 85%). All group comparisons are non-significant according to independent sample t-tests (Ps > 0.05).

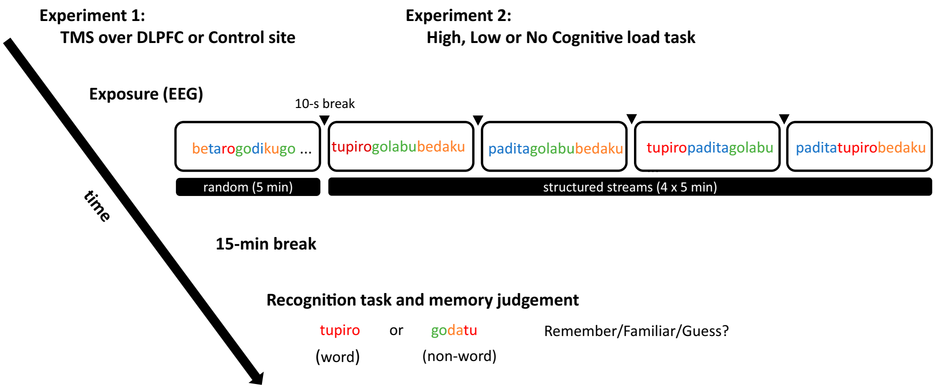

Fig. 4. Main procedure in experiments 1 and 2: In experiment 1, participants received cTBS either over DLPFC or vertex prior to exposure. In experiment 2, participants performed a dual working-memory task under HCL or LCL conditions or did not perform a task prior to exposure. Exposure started with a random stream for 5 min, after which a structured stream was presented for 20 min. A short break (10 s) was inserted every ∼5 min. EEG was recorded throughout. Participants watched a silent nature documentary during the exposure. After a 15-min break, participants completed a postexposure recognition test that examined implicit and explicit memory of the hidden words through memory judgement. During the exposure and the break, participants were not aware that the behavioral test would include segments of the syllable sequences.

(i.e., $93.3_M \pm 4.3_{SD}$ versus $81.5_M \pm 13.1_{SD}$, $t_{38} = 23.1$, $P < 0.001$, respectively). Participants under the no-load condition immediately started with the main experiment that is visualized in Fig. 4. A simple numeric self-report rating scale assessing fatigue (1: I feel no mental fatigue to 10: I feel the worst possible mental fatigue) was presented immediately before the TloadDback task and immediately after (i.e., before exposure), as a quick manipulation check for the induction of cognitive fatigue. Participants who performed no cognitive load task reported lower subjective feelings of fatigue prior to exposure (i.e., $4.2_M \pm 2.01_{SD}$) than participants who performed the cognitive load task (i.e., $5.3_M \pm 1.9_{SD}$; $P = 0.001$). However, unexpectedly, there were no reliable differences between the high- and low-load participants (i.e., $5.6_M \pm 2.1_{SD}$ versus $5.0_M \pm 2.3_{SD}$, $P = 0.21$). There were no differences in baseline subjective reports for cognitive fatigue across all groups, tested at the start of the experiment (i.e., all Ps > 0.23).

Exposure (Experiments 1 and 2). Twelve 12 unique syllables of a consonant–vowel structure were selected and structured into four novel word forms (i.e., /tu:paɪroʊ/, /goʊlɑ:bu:/, /bi:dɑ:ku:/, and /pɑ:di:tɑ:/). The individual syllables within each word form occurred at a first, second, and third position across participants so that in each stimulation group, the subjects 1 to 7 were exposed to the word forms as listed above, while subjects 8 to 14 received the word forms /paɪroʊtu:/, /lɑ:bu:goʊ/, /dɑ:ku:bi:/, and /di:tɑ:pɑ:/, and subjects 15 to 21 the word forms /roʊtu:paɪ/, /bu:goʊlɑ:/, /ku:bi:dɑ:/, and /tɑ:pɑ:di:/. This was done to minimize any stimulus-driven effects that could be caused by position preferences for syllables within a word. Across all language lists, the words were matched on average English phonotactic probability (Ps > 0.40). The 12 syllables were recorded using an online artificial speech synthesizer of a female British English voice. The audio files were edited to have a duration of 250 ms and saved with a sampling rate of 44,100 Hz using Audacity software.

Participants were informed that they would hear a continuous stream of speech sounds and were asked to listen carefully to the sounds. No information was given about the hidden structures, nor about a postexposure test on segments of the heard syllable sequences, hence language exposure was implicit. Exposure always started with a random stream, in which all 12 syllables were concatenated in a pseudorandom order without any higher-order structure; the only constraint was that syllables did not repeat and that no anagrams of the novel words appeared. In this stream, 900 syllables were presented (each syllable was repeated 75 times). After the random stream, the 20-min structured stream started, in which the speech sounds were grouped into four repeating trisyllabic words (Fig. 4). Here, the transitional probability between neighboring syllables within words was 100 and 33% between words. For instance, for subjects 1 to 3, /tu:/ in the stream is always followed by /paɪ/, while /roʊ/ could be equally followed by /goʊ/, /bi:/, or /pɑ:/.

In total, 1,200 repeating word forms (each word was repeated 300 times) and 3,600 syllables (plus the syllables from the random block) were presented. Stimulus onset asynchrony was 320 ms in experiment 1 and 310 ms in experiment 2 (this 10-ms difference between the experiments was unintended). The speech stream was presented using Presentation software (version 18.0, Neurobehavioral Systems, Inc; www.neurobs.com). The syllables were presented at a comfortable listening level for each participant through inserted earphones attached to a Dell desktop computer. During exposure, participants watched an episode of Planet Earth in silent mode (i.e., without subtitles or sound). Every ∼5 min, a short break (10 s) was inserted, which reminded the participants to attend the sounds through a visual instruction on the screen.

EEG Recording and Analyses (Experiments 1 and 2). We recorded an EEG with 27 cap-mounted electrodes (Fp1, Fp2, Fz, F3, F4, F7, F8, FC1, FC2, FC5, FC6, Cz, C3, C4, T3, T4, CP1, CP2, CP5, CP6, Pz, P3, P4, T5, T6, O1, and O2) using a TMS-compatible EEG system (TruScan Research by Deymed, Brainbox Ltd.). The nose electrode was used as reference, and an electrode attached to the forehead was used as a ground during recordings. Horizontal and vertical electrooculograms (EOGs) were recorded bipolarly with electrodes placed at the outer canthi of the eyes and above and below the right eye. Electrode impedance was kept below 10 kΩ. Signals were filtered online with a 0.1- to 1,000 Hz band-pass filter and recorded at a sampling rate of 3,000 Hz.

The EEG data analyses were performed using EEGLAB, which is an open-source toolbox operated in the MATLAB environment (60). The continuous raw data files were rereferenced to the algebraic mean of the left and right mastoid electrodes, down-sampled to 500 Hz, and filtered at 30-Hz low-pass filter. Bad channels were identified and interpolated when necessary. Mean number of interpolated channels per participant was 0.18 in experiment 1, while there were no interpolated channels in experiment 2. Periods of EEG signal containing strong muscle artifacts, electrode drifts, or technical artifacts were removed from further analysis. In experiment 1 (TMS), one participant (from the control group) was rejected due to very noisy EEG activities, while no extra participant was excluded in experiment 2. Independent component analysis was used for linear decomposition of the continuous data to remove the contributions of artifact sources (slow drifts, eye blink/movement, and muscle artifacts) on the scalp sensors. The data were segmented into 5,000-ms epochs (−2,000 to 3,000 ms relative to the onset of each word in word condition or every third syllable in the random condition). Amplitude changes exceeding ±70 μV (including the EOG channels) were removed from further analysis (less than 20% of the trials). Selective response averaging was conducted separately for each block.

We quantified neural entrainment at the syllabic and word frequencies by measuring ITC within each condition (word versus syllable/random versus structured). ITC is known as phase-locking value and is a measure of event-related phase locking. The higher the ITC value is, the higher phase coincidence across epochs is. That is, the ITC values range from 0 to 1, from purely non–phase-locked activity to strictly phase-locked activity. ITC was calculated from 0.5 to 5 Hz in 0.1-Hz steps using a continuous Morlet wavelet transformation, in which the number of cycles is increased linearly with frequency from one cycle length to obtain better frequency resolution at higher frequencies. This approach can optimize the trade-off between temporal resolution at lower frequencies and frequency resolution at high frequencies (60). Then, ITC values in each epoch were averaged.

The word presentation frequency was 1.0 and 1.1 Hz and the syllable presentation frequency was 3.1 and 3.2 Hz in experiments 1 and 2, respectively. If participants become more sensitive to the underlying word structure in the continuous speech stream, we should observe a higher ITC at the word frequency but a lower ITC at the syllable frequency during exposure to the structured sequence relative to exposure to the random sequence. In other words, if participants learn the novel words, they would show a preferential shift in the entrainment of neural oscillations to underlying words, relative to individual syllables. This can also be indexed by a simple formula, also referred to as the WLI ($\mathrm{WLI} = \mathrm{ITC}_{\text{word frequency}} / \mathrm{ITC}_{\text{syllable frequency}}$) (29). The WLI was computed across six centrofrontal midline electrodes where ITC at the word and syllable frequencies showed the strongest values (i.e., FC1, FC2, F3, F4, FC5, and Fz; supplementary documentation on ITC across electrodes is available in our open science repository, https://osf.io/dequ9/).

Forced-Choice Recognition Task (Experiments 1 and 2). After a break of ∼15 min, during which participants could remove the EEG cap and wash their hair, participants completed the two-alternative forced-choice recognition task. For each trial, a fixation cross appeared while the auditory presentation of a target three-syllable string (hidden word) and a foil three-syllable string (nonword) were presented, separated by an interstimulus interval of 1,500 ms. The nonword foils were created from the same list of 12 unique syllables that were structured into word forms. The only restriction was that the syllables within the foils never followed each other in the speech stream, not even across word boundaries. All words and nonword foils were matched on average English phonotactic probability ($Ps > 0.70$). The task was 1) to indicate which of the two strings sounded more familiar and 2) to judge on their recall decision ("I recalled from exposure" versus "It sounds familiar, but I have no clear memory" or "I guessed"). Strings that were "recalled" are referred to as confident responses, and the strings that were "familiar without memory" or "guessed" are referred to as unconfident responses. The next trial started 1,500 ms after the participant entered his or her response. The syllable strings were presented at the same rate as during the exposure. Each of the four targets and four foils were paired exhaustively to a total of 16 trials. In one-half of the trials, a target was followed by a foil, while in the other one-half, a foil was followed by a target. The order of presentation was counterbalanced across participants.

Statistical Analyses (Experiments 1 and 2). To investigate the effect of cognitive depletion on language learning, linear mixed effect analyses were performed on the EEG data (i.e., ITC values and WLI indexes), and hierarchical logistic regression analyses were performed on the behavioral data (i.e., recognition accuracy). These analyses were performed using the lme4 package (61) and the afex package (62) in R (R Development Core Team, 2011). We always strived for models including maximal random effects structure justified by the design (61, 62). In case of convergences issues (e.g., singular fits), we refitted the maximal model by first removing correlations among random slopes, after which the highest-order random slopes with the least estimated variance were removed (62). The *P* values were derived using Kenward–Roger approximations for degrees of freedom with the ANOVA function in the afex package (63). Effects coding was used for all fixed factors, except for the Group factor in experiment 2, where dummy coding was used with no cognitive load as reference level. Bonferroni correction was used for all planned tests. Cohen's d effect sizes on the model's estimates are calculated with the eff_size function from the emmeans package (64). To allow across-study comparison and facilitate secondary analyses (e.g., power calculations and meta-analyses), we also provide *t* test effect sizes in *SI Appendix*. We additionally performed one-sample *t* tests to test for above-chance performance for the unconfident and confident responses in the recognition task. One control participant in experiment 1 reported low confidence in all trials and so did not have confident responses. The stimulus materials and data files, including scripts for analysis, are available on an open science repository: https://osf.io/dequ9/).

Data Availability. Anonymized data files and scripts for analysis have been deposited to an external source (https://osf.io/dequ9/) (65).

ACKNOWLEDGMENTS. We thank Cristina Pancotto, Afrina Sallehuddin, and Ayaka Tsuchiya for help with data collection. E.H.M.S. was funded by a grant from the Research Foundation – Flanders (Grant No. 1211421N). T.D. was funded by a travel grant from the Suntory Foundation. The research costs were covered by startup funds from the School of Psychology (University of Nottingham) and the TMS and EEG equipment by a Medical Research Council fellowship (G1000566) to R.M. R.M. was also supported by Profi5 (Mind and Matter) funding from the Academy of Finland to the University of Helsinki.

1. R. A. Poldrack *et al.*, Interactive memory systems in the human brain. *Nature* **414**, 546–550 (2001).
2. M. T. Ullman, Contributions of memory circuits to language: The declarative/procedural model. *Cognition* **92**, 231–270 (2004).
3. M. H. de Vries *et al.*, Electrical stimulation of Broca's area enhances implicit learning of an artificial grammar. *J. Cogn. Neurosci.* **22**, 2427–2436 (2010).
4. J. Uddén *et al.*, The inferior frontal cortex in artificial syntax processing: An rTMS study. *Brain Res.* **1224**, 69–78 (2008).
5. M. T. Ullman, Is Broca's area part of a basal ganglia thalamocortical circuit? *Cortex* **42**, 480–485 (2006).
6. G. G. Ambrus *et al.*, When less is more: Enhanced statistical learning of non-adjacent dependencies after disruption of bilateral DLPFC. *J. Mem. Lang.* **114**, 104144 (2020).
7. G. Borragán, H. Slama, A. Destrebecqz, P. Peigneux, Cognitive fatigue facilitates procedural sequence learning. *Front. Hum. Neurosci.* **10**, 86 (2016).
8. K. Foerde, B. J. Knowlton, R. A. Poldrack, Modulation of competing memory systems by distraction. *Proc. Natl. Acad. Sci. U.S.A.* **103**, 11778–11783 (2006).
9. J. M. Galea, N. B. Albert, T. Ditye, R. C. Miall, Disruption of the dorsolateral prefrontal cortex facilitates the consolidation of procedural skills. *J. Cogn. Neurosci.* **22**, 1158–1164 (2010).
10. D. Nemeth, K. Janacsek, B. Polner, Z. A. Kovacs, Boosting human learning by hypnosis. *Cereb. Cortex* **23**, 801–805 (2013).
11. M. Virag *et al.*, Competition between frontal lobe functions and implicit sequence learning: Evidence from the long-term effects of alcohol. *Exp. Brain Res.* **233**, 2081–2089 (2015).
12. M. J. Frank, R. C. O'Reilly, T. Curran, When memory fails, intuition reigns: Midazolam enhances implicit inference in humans. *Psychol. Sci.* **17**, 700–707 (2006).
13. S. L. Thompson-Schill, M. Ramscar, E. G. Chrysikou, Cognition without control: When a little frontal lobe goes a long way. *Curr. Dir. Psychol. Sci.* **18**, 259–263 (2009).
14. A. R. Romberg, J. R. Saffran, Statistical learning and language acquisition. *Wiley Interdiscip. Rev. Cogn. Sci.* **1**, 906–914 (2010).
15. D. Kennedy, C. Norman, What don't we know? *Science* **309**, 75 (2005).
16. M. T. Ullman, "The declarative/procedural model: A neurobiological model of language learning, knowledge and use" in *The Neurobiology of Language*, G. Hickok, S. A. Small, Eds. (Elsevier, 2016), pp. 953–968.
17. K. Graf Estes, J. L. Evans, M. W. Alibali, J. R. Saffran, Can infants map meaning to newly segmented words? Statistical segmentation and word learning. *Psychol. Sci.* **18**, 254–260 (2007).
18. M. G. Gaskell, A. W. Ellis, Word learning and lexical development across the lifespan. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **364**, 3607–3615 (2009).
19. J. R. Saffran, R. N. Aslin, E. L. Newport, Statistical learning by 8-month-old infants. *Science* **274**, 1926–1928 (1996).
20. S. D. Palmer, J. Hutson, S. L. Mattys, Statistical learning for speech segmentation: Age-related changes and underlying mechanisms. *Psychol. Aging* **33**, 1035–1044 (2018).
21. J. R. Saffran, E. L. Newport, R. N. Aslin, R. A. Tunick, S. Barrueco, Incidental language learning: Listening (and learning) out of the corner of your ear. *Psychol. Sci.* **8**, 101–105 (1997).
22. R. Frost, B. C. Armstrong, M. H. Christiansen, Statistical learning research: A critical review and possible new directions. *Psychol. Bull.* **145**, 1128–1153 (2019).
23. C. M. Conway, How does the brain learn environmental structure? Ten core principles for understanding the neurocognitive mechanisms of statistical learning. *Neurosci. Biobehav. Rev.* **112**, 279–299 (2020).
24. P. Perruchet, S. Pacton, Implicit learning and statistical learning: One phenomenon, two approaches. *Trends Cogn. Sci.* **10**, 233–238 (2006).
25. T. Teinonen, V. Fellman, R. Näätänen, P. Alku, M. Huotilainen, Statistical language learning in neonates revealed by event-related brain potentials. *BMC Neurosci.* **10**, 21 (2009).
26. L. J. Batterink, K. A. Paller, P. J. Reber, Understanding the neural bases of implicit and statistical learning. *Top. Cogn. Sci.* **11**, 482–503 (2019).
27. L. J. Batterink, P. J. Reber, H. J. Neville, K. A. Paller, Implicit and explicit contributions to statistical learning. *J. Mem. Lang.* **83**, 62–78 (2015).
28. L. J. Batterink, P. J. Reber, K. A. Paller, Functional differences between statistical learning with and without explicit training. *Learn. Mem.* **22**, 544–556 (2015).
29. L. J. Batterink, K. A. Paller, Online neural monitoring of statistical learning. *Cortex* **90**, 31–45 (2017).
30. Z. Dienes, D. Berry, Implicit learning: Below the subjective threshold. *Psychon. Bull. Rev.* **4**, 3–23 (1997).
31. A. S. Finn, T. Lee, A. Kraus, C. L. Hudson Kam, When it hurts (and helps) to try: The role of effort in language learning. *PLoS One* **9**, e101806 (2014).
32. E. H. Smalle *et al.*, Can chunk size differences explain developmental changes in lexical learning? *Front. Psychol.* **6**, 1925 (2016).
33. E. H. M. Smalle, M. P. A. Page, W. Duyck, M. Edwards, A. Szmalec, Children retain implicitly learned phonological sequences better than adults: A longitudinal study. *Dev. Sci.* **21**, e12634 (2018).
34. E. H. M. Smalle, M. Panouilleres, A. Szmalec, R. Möttönen, Language learning in the adult brain: Disrupting the dorsolateral prefrontal cortex facilitates word-form learning. *Sci. Rep.* **7**, 13966 (2017).
35. E. H. M. Smalle, M. Muylle, W. Duyck, A. Szmalec, Less is more: Depleting cognitive resources enhances language learning abilities in adults. *J. Exp. Psychol. Gen.*, 10.1037/xge0001058 (2021).
36. E. L. Newport, Maturational constraints on language learning. *Cogn. Sci.* **14**, 11–28 (1990).
37. J. L. Elman, Learning and development in neural networks: The importance of starting small. *Cognition* **48**, 71–99 (1993).
38. E. L. Newport, D. Bavelier, H. J. Neville, "Critical thinking about critical periods: Perspectives on a critical period for language acquisition" in *Language, Brain, and*

Cognitive Development: Essays in Honor of Jacques Mehler, E. Dupoux, Ed. (The MIT Press, 2001), pp. 481–502.
39. E. D. Thiessen, S. Girard, L. C. Erickson, Statistical learning and the critical period: How a continuous learning mechanism can give rise to discontinuous learning. *Wiley Interdiscip. Rev. Cogn. Sci.* **7**, 276–288 (2016).
40. G. Borragán, H. Slama, M. Bartolomei, P. Peigneux, Cognitive fatigue: A time-based resource-sharing account. *Cortex* **89**, 71–84 (2017).
41. J. Bertels, A. Franco, A. Destrebecqz, How implicit is visual statistical learning? *J. Exp. Psychol. Learn. Mem. Cogn.* **38**, 1425–1431 (2012).
42. J. Breton, E. M. Robertson, Flipping the switch: Mechanisms that regulate memory consolidation. *Trends Cogn. Sci.* **18**, 629–634 (2014).
43. M. H. Christiansen, Implicit statistical learning: A tale of two literatures. *Top. Cogn. Sci.* **11**, 468–481 (2019).
44. S. Henin *et al.*, Learning hierarchical sequence representations across human cortex and hippocampus. *Sci. Adv.* **7**, eabc4530 (2021).
45. C. Benninger, P. Matthis, D. Scheffner, EEG development of healthy boys and girls. Results of a longitudinal study. *Electroencephalogr. Clin. Neurophysiol.* **57**, 1–12 (1984).
46. P. J. Marshall, Y. Bar-Haim, N. A. Fox, Development of the EEG from 5 months to 4 years of age. *Clin. Neurophysiol.* **113**, 1199–1208 (2002).
47. A. Hashemi *et al.*, Characterizing population EEG dynamics throughout adulthood. *eNeuro* **3**, 1–13 (2016).
48. J. H. Howard Jr., D. V. Howard, Aging mind and brain: Is implicit learning spared in healthy aging? *Front. Psychol.* **4**, 817 (2013).
49. D. Juhasz, D. Nemeth, K. Janacsek, Is there more room to improve? The lifespan trajectory of procedural learning and its relationship to the between- and within-group differences in average response times. *PLoS One* **14**, e0215116 (2019).
50. M. Verneau, J. van der Kamp, G. J. Savelsbergh, M. P. de Looze, Age and time effects on implicit and explicit learning. *Exp. Aging Res.* **40**, 477–511 (2014).
51. M. Muylle, E. H. Smalle, R. Hartsuiker, Rapid phonotactic constraint learning in ageing: Evidence from speech errors. *Lang. Cogn. Neurosci.*, 10.1080/23273798.2021.1897149 (2021).
52. K. McNealy, J. C. Mazziotta, M. Dapretto, Cracking the language code: Neural mechanisms underlying speech parsing. *J. Neurosci.* **26**, 7629–7639 (2006).
53. R. De Diego-Balaguer *et al.*, Striatal degeneration impairs language learning: Evidence from Huntington's disease. *Brain* **131**, 2870–2881 (2008).
54. M. F. Assaneo *et al.*, Spontaneous synchronization to speech reveals neural mechanisms facilitating language learning. *Nat. Neurosci.* **22**, 627–632 (2019).
55. W. Beam, J. J. Borckardt, S. T. Reeves, M. S. George, An efficient and accurate new method for locating the F3 position for prefrontal TMS applications. *Brain Stimul.* **2**, 50–54 (2009).
56. A. Mir-Moghtadaei *et al.*, Concordance between BeamF3 and MRI-neuronavigated target sites for repetitive transcranial magnetic stimulation of the left dorsolateral prefrontal cortex. *Brain Stimul.* **8**, 965–973 (2015).
57. J. Jung, A. Bungert, R. Bowtell, S. R. Jackson, Vertex stimulation as a control site for transcranial magnetic stimulation: A concurrent TMS/fMRI study. *Brain Stimul.* **9**, 58–64 (2016).
58. M. R. Goldsworthy, J. B. Pitcher, M. C. Ridding, A comparison of two different continuous theta burst stimulation paradigms applied to the human primary motor cortex. *Clin. Neurophysiol.* **123**, 2256–2263 (2012).
59. D. Bor, D. J. Schwartzman, A. B. Barrett, A. K. Seth, Theta-burst transcranial magnetic stimulation to the prefrontal or parietal cortex does not impair metacognitive visual awareness. *PLoS One* **12**, e0171793 (2017).
60. A. Delorme, S. Makeig, EEGLAB: An open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *J. Neurosci. Methods* **134**, 9–21 (2004).
61. D. Bates, R. Kliegl, S. Vasishth, H. Baayen, Parsimonious mixed models. arXiv [Preprint] (2015). https://arxiv.org/abs/1506.04967 (Accessed 16 June 2015).
62. H. Singmann, D. Kellen, "An introduction to mixed models for experimental psychology" in *New Methods in Cognitive Psychology*, D. H. Spieler, E. Schumacher, Eds. (Psychology Press, 2019), pp. 4–31.
63. H. Singmann *et al.*, Afex: Analysis of factorial experiments. R package version 0.13-145 (2015).
64. R. V. Lenth *et al.*, Estimated marginal means, aka least-squares means. R package version 1.6.2-1 (2021).
65. E. H. M. Smalle, R. Möttönen, Disruption of the cognitive system facilitates statistical language learning. Open Science Framework. https://osf.io/dequ9/. Deposited 14 August 2019.