

Syntactic Computation in the Human Brain: The Degree of Merger as a Key Factor

Shinri Ohta^{1,2}, Naoki Fukui^{3,4}, Kuniyoshi L. Sakai^{1,4,5*}

1 Department of Life Sciences, Graduate School of Arts and Sciences, The University of Tokyo, Komaba, Meguro-ku, Tokyo, Japan, **2** Japan Society for the Promotion of Science, Ichiban-cho, Chiyoda-ku, Tokyo, Japan, **3** Department of Linguistics, Sophia University, Kioi-cho, Chiyoda-ku, Tokyo, Japan, **4** CREST, Japan Science and Technology Agency, Goban-cho, Chiyoda-ku, Tokyo, Japan, **5** Department of Basic Science, Graduate School of Arts and Sciences, The University of Tokyo, Komaba, Meguro-ku, Tokyo, Japan

Abstract

Our goal of this study is to characterize the functions of language areas in most precise terms. Previous neuroimaging studies have reported that more complex sentences elicit larger activations in the left inferior frontal gyrus (L. F3op/F3t), although the most critical factor still remains to be identified. We hypothesize that pseudowords with grammatical particles and morphosyntactic information alone impose a construction of syntactic structures, just like normal sentences, and that “the Degree of Merger” (DoM) in recursively merged sentences parametrically modulates neural activations. Using jabberwocky sentences with distinct constructions, we fitted various parametric models of syntactic, other linguistic, and nonlinguistic factors to activations measured with functional magnetic resonance imaging. We demonstrated that the models of DoM and “DoM+number of Search (searching syntactic features)” were the best to explain activations in the L. F3op/F3t and supramarginal gyrus (L. SMG), respectively. We further introduced letter strings, which had neither lexical associations nor grammatical particles, but retained both matching orders and symbol orders of sentences. By directly contrasting jabberwocky sentences with letter strings, localized activations in L. F3op/F3t and L. SMG were indeed independent of matching orders and symbol orders. Moreover, by using dynamic causal modeling, we found that the model with an inhibitory modulatory effect for the bottom-up connectivity from L. SMG to L. F3op/F3t was the best one. For this best model, the top-down connection from L. F3op/F3t to L. SMG was significantly positive. By using diffusion-tensor imaging, we confirmed that the left dorsal pathway of the superior longitudinal and arcuate fasciculi consistently connected these regions. Lastly, we established that nonlinguistic order-related and error-related factors significantly activated the right (R.) lateral premotor cortex and R. F3op/F3t, respectively. These results indicate that the identified network of L. F3op/F3t and L. SMG subserves the calculation of DoM in recursively merged sentences.

Citation: Ohta S, Fukui N, Sakai KL (2013) Syntactic Computation in the Human Brain: The Degree of Merger as a Key Factor. *PLoS ONE* 8(2): e56230. doi:10.1371/journal.pone.0056230

Editor: William W. Lytton, SUNY Downstate MC, United States of America

Received: September 7, 2012; **Accepted:** January 8, 2013; **Published:** February 20, 2013

Copyright: © 2013 Ohta et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Funding: This research was supported by a Core Research for Evolutional Science and Technology (CREST) grant from the Japan Science and Technology Agency (JST), by a Grant-in-Aid for Scientific Research (S) (Nos. 20220005 and 22222001) from the Ministry of Education, Culture, Sports, Science, and Technology of Japan, and by a Grant-in-Aid for JSPS Fellows (No. 24-8931) from the Japan Society for the Promotion of Science (JSPS). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: sakai@mind.c.u-tokyo.ac.jp

Introduction

It is widely accepted that in human language, a sentence can be expressed by a unique tree structure with recursive branches [1,2]. Moreover, any sentence can be recursively combined within another sentence, as in e.g., “*I think that John believes that Mary assumes that...*”, and there is in principle no upper bound for the length of sentences; this property is the so-called *discrete infinity* made possible by the computational power, or engine, of the human language faculty. One possible way to elucidate the neural basis of such computational properties is to examine how the brain responds to the modulation of specified syntactic factors. An early attempt with functional magnetic resonance imaging (fMRI) has reported that activations in the language areas were modulated by noncanonical/canonical word orders and the presence/absence of lexical contents [3], in which multiple factors, including memory-related and semantic factors, could account for these activations. Therefore, we should not be content with such a general cognitive phenomenon as so-called “syntactic complexity” or “syntactic

working memory” that could involve both linguistic and non-linguistic factors. We should instead identify which minimal factor *sufficiently* explains any activation changes obtained. In addition, the size of linguistic constituents may also modulate cortical activations. A recent fMRI study has reported that the left frontal activations increased with the number of words or terminal nodes (symbols) in a phrase [4], but, as rightly pointed out by the authors, the precise phrase structures remained to be taken into account. Here we focus on different sentence constructions, and try to identify *minimal* syntactic factors associated with phrase structures, which parametrically modulate cortical responses measured with event-related fMRI.

Modern linguistics has accumulated mounting evidence that the construction of any grammatical phrases or sentences can be adequately and *minimally* explained by hierarchical syntactic structures with a set of relevant structural relations defined on such structures [5,6], leading to the postulation of the fundamental linguistic operation of *Merge* (capitalized in linguistics to indicate

formal operations), the structure-building operation, which combines two syntactic objects (words or phrases) to form a larger structure [7]. Besides Merge, we have proposed that *Search* (searching syntactic features) applies to a syntactic object already constructed by Merge, and that Search assigns relevant features to the syntactic object [8]. The total number of Merge and Search applications within an entire sentence are here simply denoted as “number of Merge” and “number of Search”, respectively. To properly measure the depth of a tree structure with a formal property of Merge and *iterativity* (recursiveness) [9], we hypothesize that “the Degree of Merger (DoM)” is a key computational concept, which can be defined as the *maximum depth* of merged subtrees (i.e., Mergers) within an entire sentence. Moreover, DoM can quantify and compare various syntactic phenomena, such as self-embedding, scrambling, wh-movement, etc. Furthermore, when Search applies to each syntactic object within hierarchical structures, the calculation of DoM plays a critical role. Indeed, from a nested sentence “[*The boy*₂ [*we*₃ [*like*₃]₂]₁ [*sings*₁]₀” (subscripts denote DoM for each node, see Figure 1A), two sentences “[*The boy ...*]₁ [*sings*]₁” and “[*we*₃ [*like*₃]₂” are obtained, where relevant features (numbers and persons here) are searched and checked between the nodes with identical DoM. Because such analyses of hierarchical structures would produce specific loads in syntactic computation, we expect that DoM and associated “number of Search” modulate neural activations. Merge would be theoretically “costless” [10,11], and thus “number of Merge” itself may not affect activations, which can be easily expected for *flat* structures (see Figure 1B).

In the present study, jabberwocky sentences that lacked lexical associations were prepared. Each sentence consisted of pseudounoun phrases (Ns) and pseudoverb phrases (Vs). We hypothesize that pseudowords with grammatical particles and morphosyntactic information alone impose a construction of syntactic structures, just like normal sentences (see **Materials and Methods, Stimuli**). Based on the nested (self-embedded), left/right-branching, and multiple-branching constructions (see Appendix S1), we introduced three basic types of sentence constructions: nested sentence (Nested), simple sentence (Simple), and conjoined sentence (Conjoined) (Figure 1A). When constructing syntactic structures like the ones shown in Figure 1A, the correspondence of each subject-verb pair is most crucial. To test that participants actually paid attention to this correspondence, we used a matching task, such that the vowel of a subject (N_i as a sample stimulus) was matched with the last vowel of the corresponding verb root (V_i as a comparison stimulus) (e.g., “zaza-ga sesaru”, underlined vowels within pseudowords). These features of vowels were only *experimentally* introduced, and this matching involved a factor of encoding (i.e., memorization of features necessary for matching). Because Vs lacked grammatical (agreement) features (e.g., number, person, gender, etc.), as in the Japanese verbs, this property of matching did not mimic agreement itself, but involved a formal association between sample and comparison stimuli. It follows that the same syntactic structures were constructed from matching and nonmatching stimuli (Tables S1 and S2), which were both well-formed, i.e., *grammatical*, in Japanese. Matching strategy (counting, for example, first and fourth stimuli for matching) was useful in solving the task, but performing the task was *not* prerequisite for constructing syntactic structures. Our matching task is different from classification tasks for symbol orders (e.g., AABB vs. ABAB, where A and B are symbols representing certain sets of stimuli), which can be solved by counting the number of each set, A or B. We further examined whether cortical activations were modulated by the length of

sentences: short (S as a subscript, four-phrase) and long (L as a subscript, six-phrase) sentences (Figure 2A).

We tested various parametric models of syntactic, other linguistic, and nonlinguistic factors (Table 1; see Appendix S1 for operational definitions), some of which were based on structure-based models (Figures S1, S2, S3). Given these factors with a limited number of experimental conditions, we wanted to narrow down the models as much as possible by adopting effective contrasts. For both short and long sentences, the estimates of “number of Merge”, as well as those of “number of case markers (-ga/-no)” and “depth of postponed symbols”, were identical among the three sentence conditions. By taking one of sentence conditions as a reference, these three factors could be eliminated from the analyses. Moreover, a reference condition should be chosen separately for each of short and long sentences, as we tested the short and long stimuli on separate days. The Conjoined condition was actually *simplest* among the three sentence conditions and thus served as an appropriate reference, because the Conjoined condition had same or less estimates than those under the Nested and Simple conditions for all factors except the numbers of Search and encoding. For brevity, a contrast with the Conjoined condition as a reference is denoted with a prime mark; e.g., [Nested_(S) – Conjoined_(S)] and [Nested_(L) – Conjoined_(L)] abbreviated as Nested'_(S) and Nested'_(L), respectively.

We further introduced letter strings, which had neither lexical associations nor grammatical particles, but retained both matching orders and symbol orders of sentences. There were two string conditions: reverse-order string (Reverse) and same-order string (Same) (Figures 1B, 2B, and Table 2). Like the sentence conditions, we used the same matching task under these string conditions, such that the first half of a string (A_i as a sample stimulus) was matched with the corresponding second half (B_i as a comparison stimulus) in the reverse or same order. These string conditions also controlled any involvement of matching strategy stated above. Between the Nested (N₂ N₁ V₁ V₂ or N₃ N₂ N₁ V₁ V₂ V₃, where each subscript denotes a matching order) and Reverse (A₂ A₁ B₁ B₂ or A₃ A₂ A₁ B₁ B₂ B₃) conditions, the curved arrows shown in Figures 1 and 2 represent the *same* matching order of sequentially presented stimuli (e.g., for N₂ N₁ V₁ V₂, the inner symbol pair of N and V is matched first, and then the outer symbol pair is matched). The symbol order was also identical among the Nested, Simple, Reverse, and Same conditions in the form of Nⁿ Vⁿ or Aⁿ Bⁿ. To control both matching orders and symbol orders, we directly compared the Nested with the Reverse, using the Simple and Same conditions as respective references (Table 2), i.e., (Nested – Simple) > (Reverse – Same). For brevity, the contrasts of [Nested – Simple] and [Reverse – Same] are denoted with a double prime mark, i.e., “Nested” and “Reverse”, respectively. Our goal with such thorough controls was to demonstrate that *purely* syntactic factors of DoM and “number of Search” actually modulate neural activations.

It has been reported that more complex sentences elicit larger activations in the pars opercularis and pars triangularis of the left inferior frontal gyrus (L. F3op/F3t) [12–19], suggesting that L. F3op/F3t is critical for syntactic processing as a grammar center [20]. On the other hand, the left angular and supramarginal gyri (L. AG/SMG) have been suggested for vocabulary knowledge or lexical processing [21,22]. To examine the functional specialization of any regions, including L. F3op/F3t and L. AG/SMG, in an unbiased manner, we adopted whole-brain analyses [23]. We also performed effective connectivity analyses by using dynamic causal modeling (DCM) [24] to examine the functional integration of identified regions. To provide empirical backup for the connection derived from DCM, we checked the anatomical plausibility of the

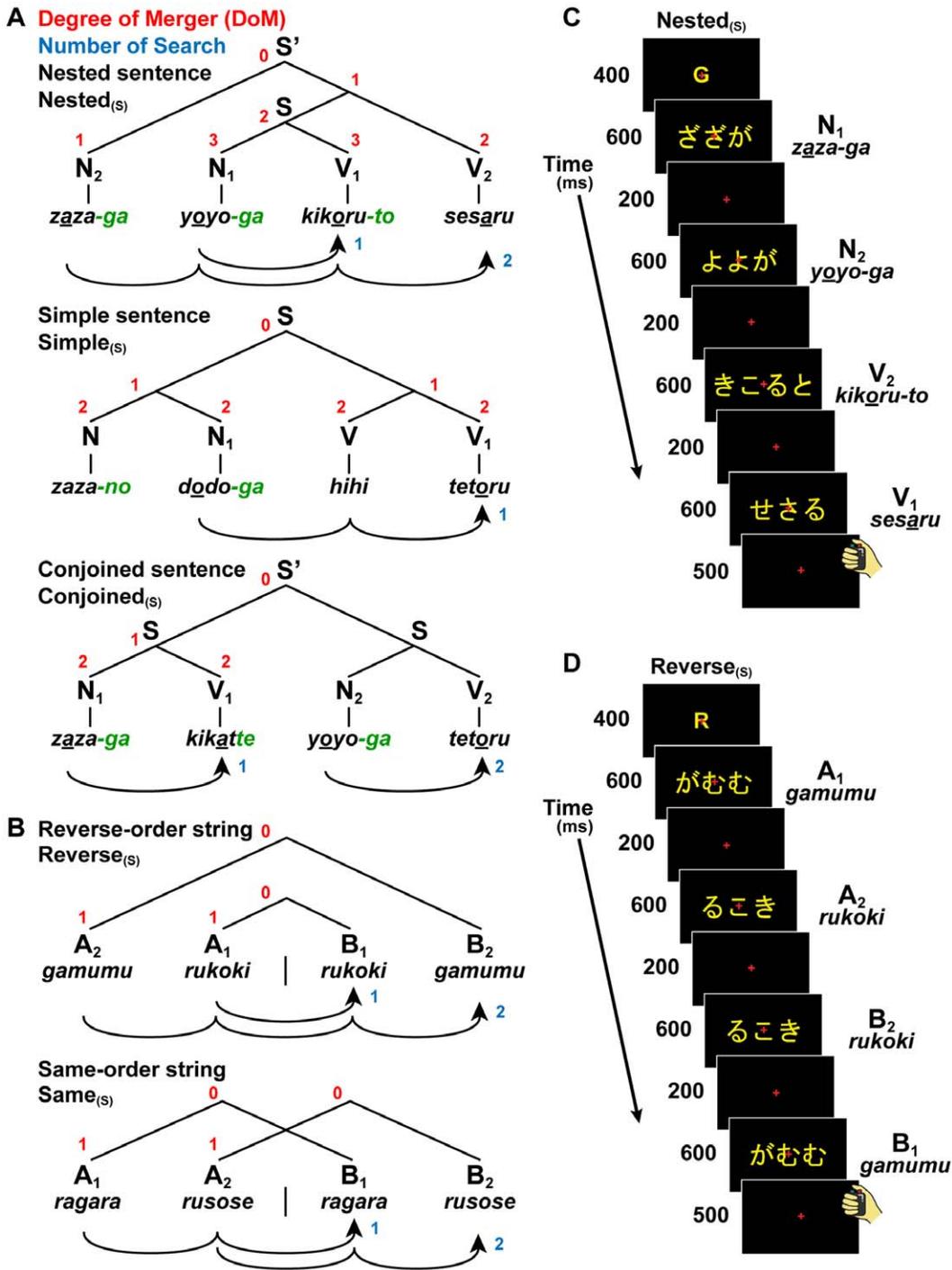


Figure 1. A paradigm for testing jabberwocky sentences and letter strings. Examples of short [(S) as a subscript] matching stimuli are shown here with the Romanization system, but actual stimuli were presented in hiragana without hyphen (see C and D). (A) Three sentence conditions with short stimuli: Nested_(S), Simple_(S), and Conjoined_(S). Based on contemporary linguistics [2], each diagram represents a unique tree structure of each sentence (S and S') constructed from Ns and Vs. For the Nested_(S), a sentence (S) at the lowest hierarchical level was nested into an entire sentence (S'), similar to "Taro-ga Hanako-ga utau-to omou" ("Taro thinks that Hanako sings"). For the Simple_(S), a simple sentence was constructed by adding the same number of left/right branches to both Ns and Vs, similar to "Taro-no ani-ga tabe hajimeru" ("Taro's brother starts eating"). For the Conjoined_(S), an entire sentence (S') was constructed by conjoining two sentences, similar to "Taro-ga utatte Hanako-ga odoru" ("Taro sings, and Hanako dances"). The digits shown in red and blue denote DoM for each node and "number of Search", respectively (see Table 1). The curved arrows denote the matching of sequentially presented stimuli. (B) Two string conditions with short stimuli: Reverse_(S) and Same_(S). Each letter string was formed by jumbling letters of either N or V. (C and D) Examples of stimulus presentation. Here, examples of matching stimuli are shown in hiragana for the Nested_(S) and Reverse_(S). Between the Nested_(S) and Reverse_(S), both of the symbol orders (the order of Ns, Vs, As, and Bs) and matching orders (denoted by subscripts) were identical. doi:10.1371/journal.pone.0056230.g001

Table 1. Estimates of various factors to account for activations under the sentence conditions.

Syntactic factors							
Factor	Nested _(L)	Nested _(S)	Simple _(L)	Simple _(S)	Conjoined _(L)	Conjoined _(S)	
Degree of Merger (DoM)	5	3	3	2	2	2	
No. of Search	3	2	2	1	3	2	
No. of Merge	5	3	5	3	5	3	
	Nested' _(L)	Nested' _(S)	Simple' _(L)	Simple' _(S)			
DoM	3	1	1	0			
DoM + No. of Search	3	1	0	-1			
No. of Search	0	0	-1	-1			
No. of Merge	0	0	0	0			
Other linguistic factors							
Factor	Nested _(L)	Nested _(S)	Simple _(L)	Simple _(S)	Conjoined _(L)	Conjoined _(S)	
No. of case markers (-ga/-no)	3	2	3	2	3	2	
No. of tense markers (-ru/-ta)	3	2	1	1	1	1	
Degree of nesting	2	1	1	1	1	1	
Degree of self-embedding	2	1	1	0	1	0	
No. of nodes	11	7	11	7	10	7	
Depth of postponed symbols	3	2	3	2	3	2	
Integration costs	5	3	3	2	1	1	
Storage costs	3	2	2	2	1	1	
Syntactic interference	2	1	0	0	0	0	
Positional similarity	3	2	2	0	0	0	
	Nested' _(L)	Nested' _(S)	Simple' _(L)	Simple' _(S)			
No. of case markers (-ga/-no)	0	0	0	0			
No. of tense markers (-ru/-ta)	2	1	0	0			
Degree of nesting	1	0	0	0			
Degree of self-embedding	1	1	0	0			
No. of nodes	1	0	1	0			
Depth of postponed symbols	0	0	0	0			
Integration costs	4	2	2	1			
Storage costs	2	1	1	1			
Syntactic interference	2	1	0	0			
Positional similarity	3	2	2	0			
Nonlinguistic factors							
Factor	Nested _(L)	Nested _(S)	Simple _(L)	Simple _(S)	Conjoined _(L)	Conjoined _(S)	
Memory span	4	2	2	1	0	0	
Counting	2	1	2	1	0	0	
No. of encoding	6	4	3	2	6	4	
	Nested' _(L)	Nested' _(S)	Simple' _(L)	Simple' _(S)			
Memory span	4	2	2	1			
Counting	2	1	2	1			
No. of encoding	0	0	-3	-2			
Memory span + counting	6	3	4	2			
Memory span + No. of encoding	4	2	-1	-1			

We define the estimate of a factor as the largest value that the factor can variably take within an entire sentence. For each factor, its unit load should be invariable among all sentence conditions, making an independent subtraction between estimates of the same factor possible. Separately for long and short sentences, estimates under the Conjoined condition as a reference were subtracted from those under the other Nested and Simple conditions (e.g., the [Nested_(L) - Conjoined_(L)] contrast abbreviated as Nested'_(L); e.g., DoM for Nested'_(L), 5-2=3). Excluding "number of Merge", the estimates of which were null for all four contrasts, we regarded "DoM+number of Search" (i.e., adding the estimates of two factors) as an additional factor. Among the nonlinguistic factors, "memory span+counting" and "memory span+number of encoding" were regarded as additional factors, because they were temporal order-related and memory-related factors, respectively.
doi:10.1371/journal.pone.0056230.t001

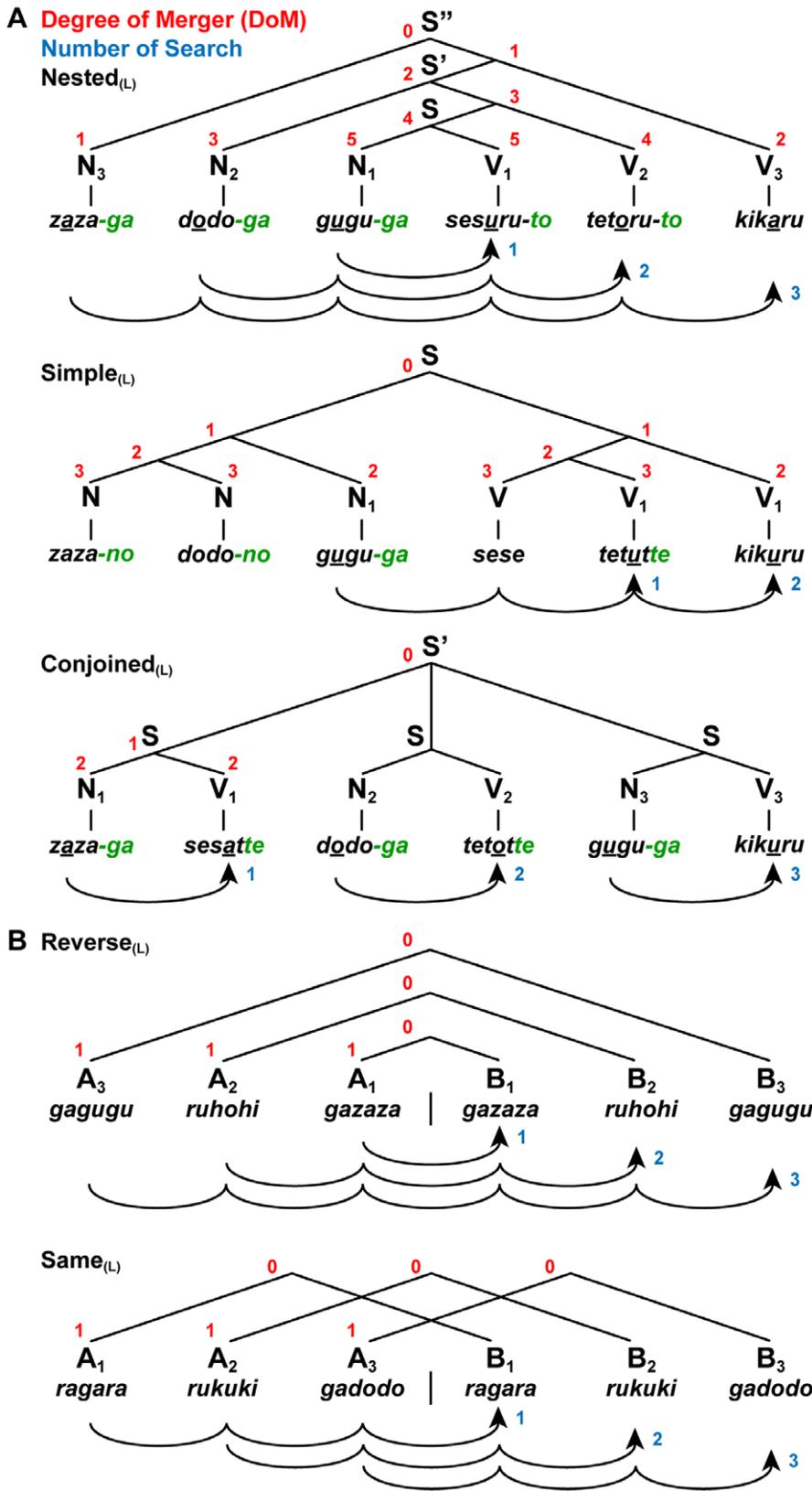


Figure 2. Examples of long matching stimuli. (A) Three sentence conditions with long [(L) as a subscript] stimuli: Nested_(L), Simple_(L), and Conjoined_(L). (B) Two string conditions with long stimuli: Reverse_(L) and Same_(L). See Appendix S2 for further explanation. doi:10.1371/journal.pone.0056230.g002

Table 2. Estimates of nonlinguistic and syntactic factors to account for activations.

Nonlinguistic factors									
Factor	Nested _(L)	Nested _(S)	Simple _(L)	Simple _(S)	Reverse _(L)	Reverse _(S)	Same _(L)	Same _(S)	
Memory span	4	2	2	1	4	2	2	1	
Counting	2	1	2	1	2	1	2	1	
No. of encoding	6	4	3	2	6	4	6	4	
	Nested		Simple		Reverse		Same		
Memory span	6		3		6		3		
Counting	3		3		3		3		
No. of encoding	10		5		10		10		
	Nested''				Reverse''				
Memory span	3				3				
Counting	0				0				
No. of encoding	5				0				
Syntactic factors									
Factor	Nested''				Reverse''				
DoM	3				0				
DoM+No. of Search	5				0				

For Nested, Simple, Reverse, and Same, the estimates for short and long stimuli were added together, because each factor's unit load would be invariable between short and long stimuli under each of the sentence and string conditions. Because the matching orders or symbol orders were identical between the Nested and Reverse conditions, the unit load of memory span or counting was invariable between the Nested and Reverse conditions, which was also invariable between the Reverse and Same conditions, thus invariable among the Nested, Simple, Reverse, and Same conditions. For brevity, the contrasts of [Nested – Simple] and [Reverse – Same] are denoted with a double prime mark, i.e., "Nested'' and Reverse'', respectively. Note that the estimates of memory span in Nested'' and Reverse'' also became identical, and that the Reverse'' contrast makes the listed estimates null, except memory span. The last two syntactic factors, whose models were best in Table 5, consistently accounted for the results of Figure 4F. All estimates of the other factors unlisted here were null in Reverse'', which cannot account for the results of Figure 5C and 5D. doi:10.1371/journal.pone.0056230.t002

network with diffusion-tensor imaging (DTI). According to recent DTI studies, there have been controversial issues as regards the functional roles of two different pathways for syntax, semantics, and phonology: dorsal tracts of the superior longitudinal and arcuate fasciculi (SLF/AF), as well as ventral tracts of the middle longitudinal fasciculus (MdLF) and extreme capsule (EmC); both pathways connect the inferior frontal and superior/middle temporal areas [25–28]. Our present study would elucidate the most crucial network and pathway for syntactic computation.

Materials and Methods

Participants

Eighteen native Japanese speakers (all males, aged 19–25 years), who had not majored in linguistics, participated in an fMRI experiment. Additional 15 participants (14 males, aged 19–40 years) were tested in a DTI experiment. All participants in the fMRI and DTI experiments were healthy and right-handed (laterality quotients: 11–100), according to the Edinburgh inventory [29]. Prior to participation in the study, written informed consent was obtained from each participant after the nature and possible consequences of the studies were explained. Approval for the experiments was obtained from the institutional review board of the University of Tokyo, Komaba.

Stimuli

Each visual stimulus consisted of two to five yellow letters in hiragana (Figure 1C and 1D). The stimuli were visually presented against a dark background through an eyeglass-like MRI-compatible display (resolution, 800×600; VisuaStim XGA; Resonance Technology Inc., Northridge, CA). The visual stimuli were always presented at the center of the monitor. At the

initiation of every trial of the Nested, Simple, and Conjoined, the cue "G" (for grammar conditions with all grammatical sentences) was shown for 400 ms. The cue "R" (for reverse orders) was shown for the Reverse, and "M" (for memorizing orders) for the Same. Four (short) or six (long) stimuli were each sequentially presented to the participants for 600 ms, with an interstimulus interval of 200 ms, leading to 4.5 s and 6 s trials for the short and long stimuli, respectively. For fixation, a red cross was always displayed at the center of the monitor. During fMRI experiments, stimulus presentation, as well as acquisition of responses and reaction times (RTs), was controlled using the LabVIEW software and interface (National Instruments, Austin, TX).

Under the sentence conditions, jabberwocky sentences consisting of pseudonoun phrases and pseudoverb phrases alone were presented in a phrase-by-phrase manner to the participants. We made six pseudonouns by repeating the same syllables with voiced consonants and any one of /a/, /u/, or /o/: *rara*, *zaza*, *mumu*, *gugu*, *yoyo*, and *dodo*. We also made four pseudoverb roots by repeating the same syllables with voiceless consonants and either /i/ or /e/: *kiki*, *hihi*, *sese*, and *tete*. The transitions between consecutive phrases or sentences were thoroughly randomized. Nonmatching stimuli included at least one odd vowel of V_i as a matching error (Tables S1 and S2). All matching and nonmatching stimuli were phonotactically legal, but lacked lexical associations in Japanese. There were 10 conditions (Figures 1 and 2); we prepared a set of 36 sentences for each of sentence conditions, and a set of 36 letter strings for each of string conditions. Each set consisted of 18 matching and 18 nonmatching stimuli. See Appendix S2 for detailed information about the stimuli.

We used only three kinds of grammatical particles, which represent *canonical* (i.e., in a prototypical use) case markings and syntactic information in Japanese: *-ga*, a nominative case marker; -

no, a genitive case marker; and *-to*, a complementizer. In all jabberwocky sentences, the distinction between Ns and Vs was clear without memorizing pseudowords, because Ns, but not Vs, ended with either *-ga* or *-no*; only nouns and pronouns precede case markers in Japanese (e.g., “*momo-ga minoru*” and “*momo-no iro*”: “*the peach ripens*” and “*the peach’s color*”; real phrases will be translated hereafter). Moreover, Vs took a nonpast-tense form (*-ru*), past-tense form (*-ta*), or gerundive form (*-te*), following morpho-syntactic and phonological features of Japanese verbs [30]; Vs ended with *-to* and *-te* introduced *that*-clauses and *and*-conjunctives, respectively (see examples in Figure 1 legend). Including the first verb of a compound verb in an adverbial form (e.g., “*hishi*” and “*sese*”), all Ns and Vs with *-ga*, *-no*, *-to*, and *-te* endings (green letters in Figures 1A and 2A) were associated with Merge applications to connect multiple nouns/verbs or sentences, amounting to “number of Merge”.

Under the string conditions, stimuli were presented in the reverse order for the Reverse, whereas they were in the same order for the Same, as regards the first and second halves of a string (Figures 1B and 2B). Each letter string was formed by jumbling letters of either N or V, which had no lexical associations. For the Reverse and Same, there was actually no path connecting the nonterminal nodes of symbol pairs (e.g., $A_1 B_1$ and $A_2 B_2$), as there was *no* Merge application to connect the multiple pairs. The letter strings lacked *-ga*, *-no*, *-to*, or *-te* endings, and their flat constructions were determined by the cue of “R” or “M” alone. We estimated the syntactic factors for the letter strings, but all estimates of these factors were null in Reverse” (see Table 2).

Task

For each trial of a matching task under the sentence conditions or string conditions, the participants judged whether or not all pairs of the sample stimulus (N or A) and comparison stimulus (V or B) were matched, and responded by pressing one of two buttons (right for matching, and left for nonmatching) after the last stimulus appeared (Figure 1C and 1D). The accuracy and RTs were collected until 500 ms after the last stimulus disappeared. No feedback on each trial’s performance was given to any participant. See Appendix S3 for task instructions and training procedures.

For the Nested, an entire sentence was constructed by nesting sentences in the form of $[N_2[N_1 V_1]V_2]$ or $[N_3[N_2[N_1 V_1]V_2]V_3]$, where $[N_i V_i]$ represents a subject-verb pair of a sentence (Figures 1A and 2A). In head-last languages, the key element (the “head”) that determines the properties of a phrase is placed at the end of the phrase. Because Japanese is a head-last, and hence an SOV (verb-final) language, a main verb is placed after a subordinate clause. Therefore, Japanese sentences naturally yield nested structures of $N^n V^n$ without having to employ, as in English, object-relative clauses (e.g., “*The boy who_i we like_t sings_s*”), which require “movement” of an object (i.e., with more Merge applications) leaving behind a “trace” (t_i). For the Simple, a simple sentence was constructed by adding the same number of left/right branches to both Ns and Vs. The last noun (i.e., head) in the branches of Ns made a subject-verb pair with the last verb (i.e., head) of a compound verb. Each simple sentence thus took the form of $[(NN_1) (VV_1)]$, etc. For the Conjoined, an entire sentence was constructed by conjoining sentences in the form of $[N_1 V_1][N_2 V_2]$ or $[N_1 V_1][N_2 V_2][N_3 V_3]$.

In a single run of 60 trials for the short stimuli, there were 10 trials each for the sentence conditions (the Nested_(S), Simple_(S), and Conjoined_(S)), and 15 trials each for the string conditions (the Reverse_(S) and Same_(S)). Each trial was alternately a sentence condition and a string condition. If the sentence and string sequences were separated, the order of the Nested, Simple, and

Conjoined was pseudo-randomized without repetition, and the order of the Reverse and Same was counterbalanced as Same-Reverse-Reverse-Same-... or Reverse-Same-Same-Reverse-... In a single run of 50 trials for the long stimuli, there were 10 trials each for the sentence conditions (the Nested_(L), Simple_(L), and Conjoined_(L)) and the string conditions (the Reverse_(L) and Same_(L)), in the order of string-sentence-string-sentence-sentence-string-... With a maximum of nine runs, the same sentence stimulus appeared no more than three times for each participant.

MRI Data Acquisition

Depending on the time of experiments, the fMRI scans were conducted on a 1.5 T scanner (Stratis II, Premium; Hitachi Medical Corporation, Tokyo, Japan) with a bird-cage head coil, and the DTI scans were conducted on a 3.0 T scanner (Signa HDxt; GE Healthcare, Milwaukee, WI) with an 8-channel phased-array head coil. For the fMRI, we scanned 26 axial slices that were 3-mm thick with a 1-mm gap, covering from $z = -40$ to 63 mm from the anterior to posterior commissure (AC-PC) line, with a gradient-echo echo-planar imaging (EPI) sequence [repetition time (TR) = 3 s, echo time (TE) = 51 ms, flip angle (FA) = 90°, field of view (FOV) = 192 × 192 mm², resolution = 3 × 3 mm²]. In a single scanning run, we obtained 92 volumes for the short stimuli and 101 volumes for the long stimuli following three dummy images, which allowed for the rise of the MR signals. For each participant, five to nine runs for each of the short and long stimuli were tested, and four to nine runs without head movement were used for analyses. After completion of the fMRI session, high-resolution T1-weighted images of the whole brain (145 axial slices, 1 × 1 × 1 mm³) were acquired from all participants with a radio frequency spoiled steady-state acquisition with a rewind gradient echo sequence (TR = 30 ms, TE = 8 ms, FA = 60°, FOV = 256 × 256 mm²).

For the DTI, we scanned 50 axial slices that were 3-mm thick without gap, covering from $z = -60$ to 90 mm from the AC-PC line, with a diffusion-weighted spin-echo EPI sequence (b-value = 1,000 s/mm², TR = 15 s, TE = 87 ms, FOV = 256 × 256 mm², resolution = 2 × 2 mm², number of excitations = 2). A single image without diffusion-weighting (b0) was initially acquired, and then diffusion-weighting was isotropically distributed along 60 diffusion-encoding gradient directions. After completion of the DTI sessions, high-resolution T1-weighted images of the whole brain (192 axial slices, 1 × 1 × 1 mm³) were acquired from all participants with a fast spoiled gradient recalled acquisition in the steady state sequence (TR = 10 ms, TE = 4 ms, FA = 25°, FOV = 256 × 256 mm²). See Appendix S4 for MRI data analyses.

Results

Condition and Length Effects on the Accuracy/RTs

The accuracy data, as well as RTs measured from the onset of the last stimulus, are shown in Figure 3. The high accuracy under both sentence and string conditions indicated the participants’ reliable and consistent judgments on the matching task. A two-way repeated-measures analysis of variance (rANOVA) with the condition [Nested, Simple, Conjoined, Reverse, Simple] × length [Long, Short] for the accuracy showed a significant main effect of condition [$F(4, 68) = 15, P < 0.0001$] and an interaction of condition by length [$F(4, 68) = 12, P < 0.0001$], but a main effect of length was not significant [$F(4, 68) = 3.8, P = 0.07$]. The RTs also showed a significant main effect of condition [$F(4, 68) = 43, P < 0.0001$] and an interaction of condition by length [$F(4, 68) = 13, P < 0.0001$], but a main effect of length was not

significant [$F(4, 68) = 1.1, P = 0.30$]. Post-hoc paired t -tests among all conditions (significance level at $\alpha = 0.005$, Bonferroni corrected) showed that the accuracy for the Nested was significantly lower than that under the other conditions including the Reverse ($P < 0.0001$). This result indicates that the Nested was the most demanding condition, which cannot be explained by the *non-linguistic* factors we examined (cf. the same estimates for the Nested and Reverse in Table 2, as well as its notes). On the other hand, post-hoc paired t -tests showed that the RTs under each sentence condition were significantly longer than those under each string condition ($P < 0.0001$). This difficulty was not in the task itself, but in vowel extraction; the sentence conditions, but not the string conditions, involved vowel extraction from the second syllable of V_i presented in hiragana, especially for the last V_i that were directly linked with RTs (Figure 1C and 1D). The load of vowel extraction would become also larger for the *short* stimuli, as we tested the short and long stimuli on separate days in the order short, then long. Indeed, the accuracy for the Conjoined_(S) was significantly lower than that for the Conjoined_(L) [$t(17) = 3.1, P = 0.006$] (significance level at $\alpha = 0.01$, Bonferroni corrected), and the RTs for the Conjoined_(S) were significantly longer than those for the Conjoined_(L) [$t(17) = 2.8, P = 0.01$], probably reflecting associated effects for novices. For the Conjoined, length effects were apparently absent, and the estimates of both memory span and counting, which were associated with length effects, were indeed null for the Conjoined alone (Table 1). In the present study, we mainly analyzed activations that would show length effects (i.e., Long>Short), excluding the involvement of vowel extraction or effects for novices. Moreover, we used the Conjoined condition, which showed such effects most strongly, as a reference for both Nested and Simple conditions. Therefore, we can safely conclude that any elicited effects did not directly relate to the task.

Under the string conditions, the accuracy for the long stimuli was significantly lower than that for the short stimuli ($P < 0.001$), indicating length effects. For the Nested and Simple conditions, in contrast, the effects for novices and length would have been cancelled out, as neither the accuracy nor RTs differed significantly between the short and long stimuli ($P > 0.05$). Under the string conditions, the accuracy was more sensitive than the RTs.

Functional Evidence of Syntactic Computation in Language Areas

We examined brain activation under the sentence conditions, in particular focusing on selective activations for the most-demanding Nested condition. In a two-way analysis of covariance (ANCOVA) with the condition [Nested', Simple'] \times length [Long, Short], the main effect of condition, i.e., Nested' $>$ Simple' while combining Long and Short, resulted in left-dominant activation, especially in L. F3op/F3t, left lateral premotor cortex and F3op (L. LPMC/F3op), and L. SMG (Figure 4A and Table 3). Other significantly activated regions were the right (R.) F3op/F3t, R. LPMC, anterior cingulate cortex (ACC), and R. SMG. The main effect of length, i.e., Long>Short while combining Nested' and Simple', also showed significant activations in the same regions, while there were more significant voxels in the right hemisphere (Figure 4B). Therefore, length effects alone cannot account for the consistent activation in these regions. An interaction of condition by length did not show any significant activation.

To further narrow down candidate regions, we tested Nested'_(L)>Simple'_(S), which reflected both main effects, and found significant activation in L. F3op/F3t and L. SMG (Figure 4C and Table 4). The data used for selecting these regions and those for subsequent analyses were not independent, which

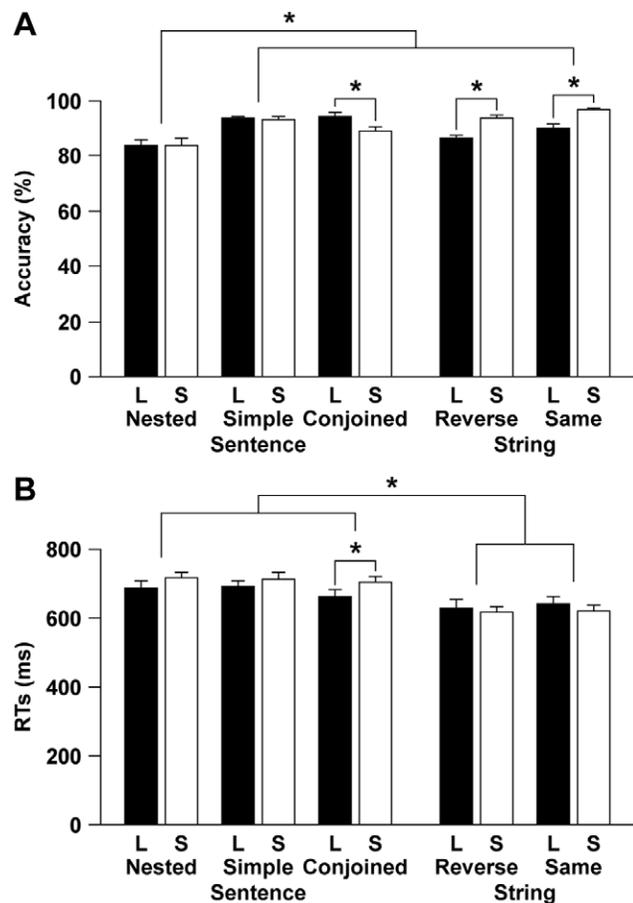


Figure 3. Condition and length effects on the accuracy/RTs. (A) The accuracy (mean \pm SEM) for long (L) and short (S) stimuli, denoted by filled and open bars, respectively. Asterisks indicate the significance level at corrected $P < 0.05$ (paired t -tests). (B) RTs from the onset of the last stimulus.

doi:10.1371/journal.pone.0056230.g003

might cause a selection bias [31]. Among the four contrasts, however, Nested'_(L) and Simple'_(S) yielded two extremes of the estimates of most factors, without apparent bias among the factors (see Table 1). In addition to both main effects, the percent signal changes in L. F3op/F3t and L. SMG (Figure 4D), averaged across significant voxels, showed a significant length effect within each of Nested' and Simple' (paired t -test, $P < 0.02$; significance level at $\alpha = 0.025$, Bonferroni corrected). Because we used appropriate references of the Conjoined_(L) and Conjoined_(S), we examined whether likewise *subtracted* estimates of each factor (e.g., DoM for Nested'_(L); see Table 1) directly explained parametric modulation of activations in the four contrasts of Nested'_(L), Nested'_(S), Simple'_(L), and Simple'_(S). The signal changes in L. F3op/F3t and L. SMG indeed correlated exactly in a step-wise manner with the parametric models of DoM [3, 1, 1, 0] and “DoM+number of Search” [3, 1, 0, -1], respectively.

Next we examined how well activations in L. F3op/F3t and L. SMG correlated with DoM and other factors. All contrasts of Nested'_(L), etc. predicted that activations should be exactly zero when a factor produced no effect or load relative to the Conjoined. We thus adopted a no-intercept model, in which percent signal changes of each region were fitted with a single (thus minimal) scale parameter to a model of each factor using its subtracted estimates. For the four contrasts, a least-squares method was used

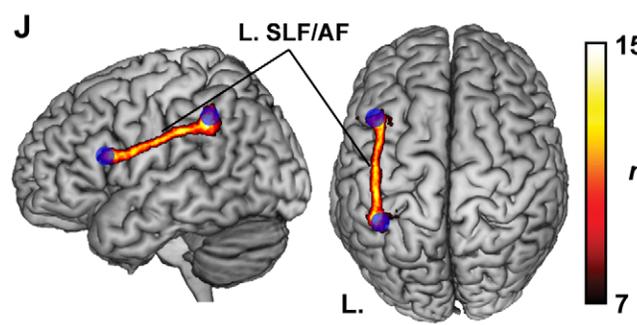
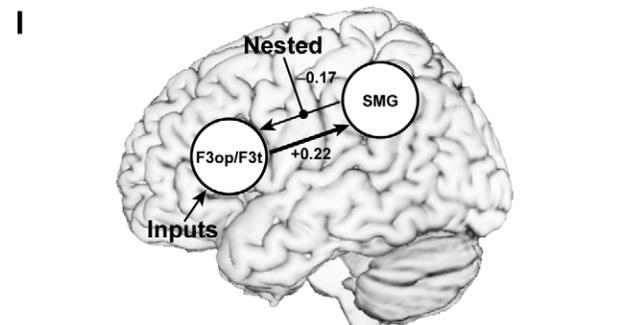
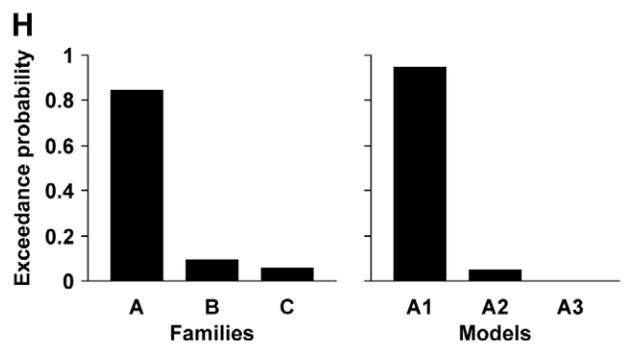
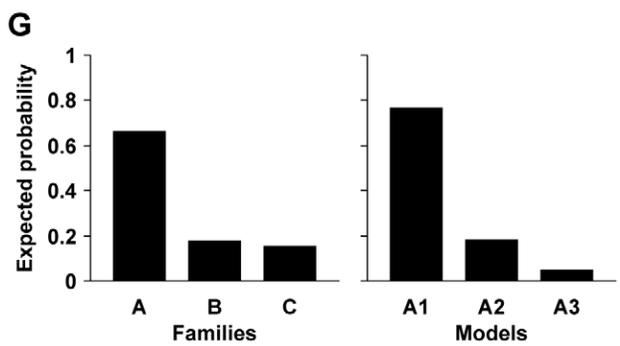
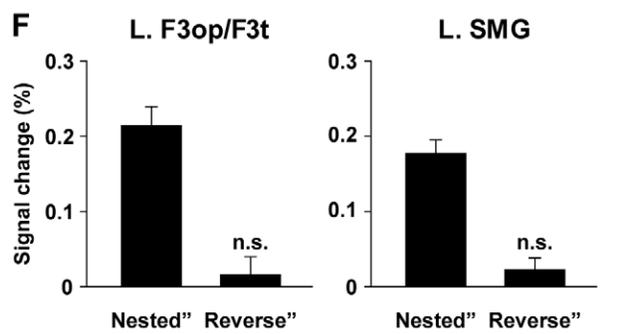
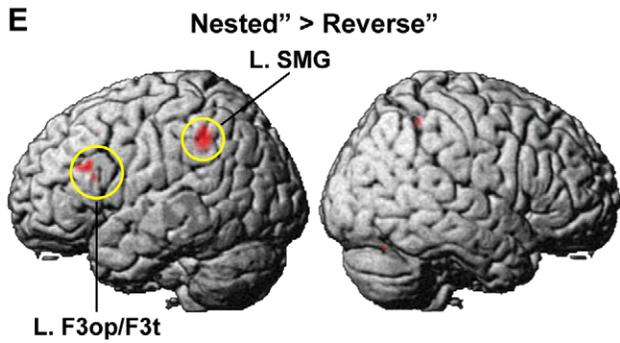
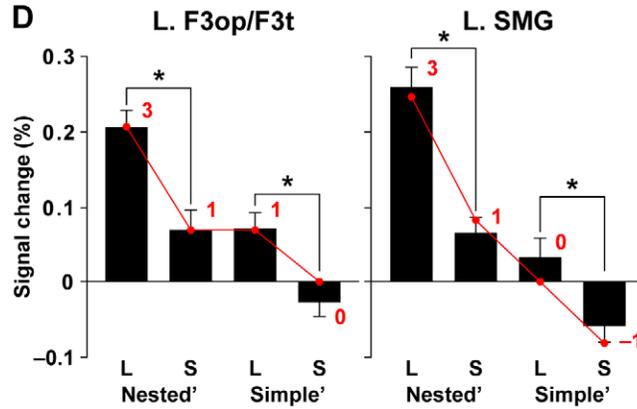
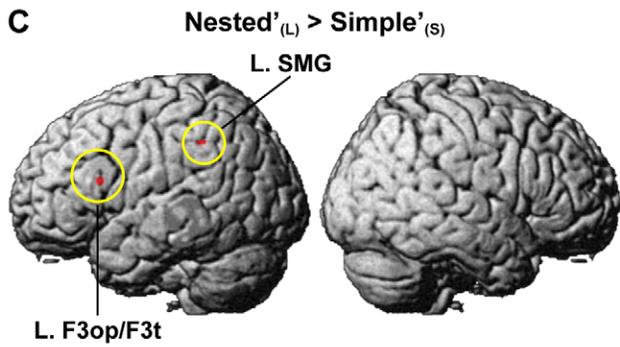
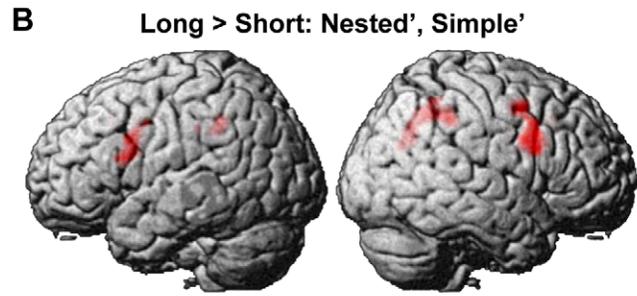
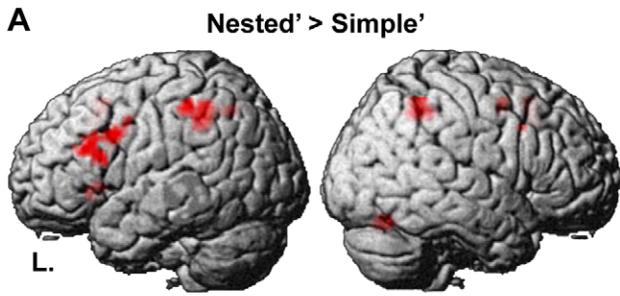


Figure 4. Functional and anatomical evidence of syntactic computation in language areas. For (A) and (B), we used a two-way ANCOVA with condition×length; for (C) and (E), a one-way ANCOVA was used. Activations were projected onto the left (L) and right lateral surfaces of a standard brain. See Tables 3 and 4 for their stereotactic coordinates. (A) Regions identified by the main effect of condition, i.e., Nested’>Simple’ (Nested’ and Simple’ denote [Nested – Conjoined] and [Simple – Conjoined], respectively). (B) Regions identified by the main effect of length, i.e., Long>Short while combining Nested’ and Simple’. (C) Regions identified by Nested’_(L)>Simple’_(S), which reflected both main effects. (D) Percent signal changes for Nested’ and Simple’, averaged across L. F3op/F3t and L. SMG in (C) (mean ± SEM). Overlaid red dots and lines denote the values fitted with the estimates (digits in red) for the best models: DoM for L. F3op/F3t and “DoM+number of Search” for L. SMG. (E) Regions identified by Nested’>Reverse” (Nested” and Reverse” denote [Nested – Simple] and [Reverse – Same], respectively). (F) Percent signal changes for Nested” and Reverse”, averaged across the L. F3op/F3t and L. SMG in (E). (G–I) The results of DCM, testing effective connectivity between L. F3op/F3t and L. SMG (see Figure S4). Bar graphs show expected probabilities (G) and exceedance probabilities (H) for each modulatory family and for the input models of the winning family A. The best model A1 (I) included a significant intrinsic connection (a thick line). (J) Anatomical connectivity between L. F3op/F3t and L. SMG revealed by DTI. The population probability map is shown on the left lateral and dorsal surfaces of a standard brain with maximum intensity projection. Blue spheres represent seed regions of L. F3op/F3t and L. SMG.

doi:10.1371/journal.pone.0056230.g004

to minimize residual sum of squares (RSS) for the four fitted values (i.e., four estimates multiplied by the fitting scale) against corresponding signal changes averaged across participants (Table 5). Among a number of parametric models tested, the model of DoM for L. F3op/F3t, as well as that of “DoM+number of Search” for L. SMG, produced by far the least RSS (≤ 0.0020) and largest coefficient of determination ($r^2 \geq 0.97$). Goodness of fit was further evaluated for each model by using a one-sample *t*-test (significance level at $\alpha = 0.0125$, Bonferroni corrected) between the fitted value for each contrast and individual activations. The model of DoM for L. F3op/F3t, as well as that of “DoM+number of Search” for L. SMG, produced no significant deviation for the four contrasts (one-sample *t*-test, $P \geq 0.17$). For L. SMG, the second-best model was DoM (RSS = 0.0063, $r^2 = 0.92$, and its

smallest $P = 0.013$ was marginal). To further take account of interindividual variability, we fitted “linear mixed-effects models” with individual activations (Table 5), and found that the models of DoM and “DoM+number of Search” were by far more likely for L. F3op/F3t and L. SMG, respectively.

Next, we examined whether the selective activation in these regions was replicated even after controlling both matching orders and symbol orders (e.g., $N_2 N_1 V_1 V_2$ and $A_2 A_1 B_1 B_2$) between the Nested and Reverse, i.e., in Nested”>Reverse” combining the short and long stimuli. This contrast indeed resulted in significant activation in L. F3op/F3t and L. SMG (Figure 4E and Table 4). In both regions, the signal changes in Reverse” were not significantly different from 0 (one-sample *t*-test, $P > 0.1$) (Figure 4F). Moreover, the models of DoM and “DoM+number of Search” were also

Table 3. Regions related to the sentence conditions.

Contrast	Brain region	BA	Side	x	y	z	Z Value	Voxels
Main effect of condition, Nested’>Simple’	F3op/F3t	44/45	L	-51	27	24	5.6	109
	LPMC/F3op	6/44	L	-48	9	30	5.4	*
	F3op/F3t	44/45	R	54	15	36	4.9	2
	LPMC	6	R	33	3	51	5.3	12
	Insula	-	L	-30	24	-3	5.7	20
	ACC	6/8/32	M	-3	18	48	6.1	45
	SMG	40	L	-54	-33	48	5.3	101
				-39	-42	39	5.9	*
				42	-48	54	5.3	64
		AG/SMG	39/40	L	-30	-60	48	4.9
	Cerebellum, lobule VI	-	R	27	-69	-21	5.6	26
Main effect of length, Long>Short: Nested’, Simple’	F3op/F3t	44/45	L	-48	12	18	5.9	63
	LPMC/F3op	6/44	L	-48	3	39	4.7	3
			R	48	6	30	6.0	129
	F3op/F3t	44/45	R	54	12	30	5.8	*
	LPMC	6	R	30	0	48	5.9	*
	ACC	6/8/32	M	0	27	39	4.9	9
	SMG	40	L	-57	-30	36	4.7	1
				-36	-45	39	5.3	26
				42	-42	42	5.5	116
		AG/SMG	39/40	R	33	-63	27	5.4

Stereotactic coordinates (x, y, z) in the Montreal Neurological Institute (MNI) space (mm) are shown for each activation peak of Z values (corrected $P < 0.05$). BA, Brodmann’s area; L, left hemisphere; R, right hemisphere; M, medial; F3op/F3t, pars opercularis and pars triangularis of the inferior frontal gyrus; LPMC, lateral premotor cortex; ACC, anterior cingulate cortex; SMG, supramarginal gyrus; AG, angular gyrus. The region with an asterisk is included within the same cluster shown one row above.

doi:10.1371/journal.pone.0056230.t003

Table 4. Regions related to the sentence conditions and/or string conditions.

Contrast	Brain region	BA	Side	x	y	z	Z Value	Voxels
Nested _(L) >Simple _(S)	F3op/F3t	44/45	L	-45	18	18	4.8	1
	SMG	40	L	-42	-45	42	4.8	2
Nested">Reverse"	F3op/F3t	44/45	L	-51	24	24	5.8	5
	ACC	6/8/32	M	-3	18	45	5.2	1
	SMG	40	L	-39	-45	42	5.7	27
			R	39	-48	54	4.9	2
	Cerebellum, lobule VI	-	R	27	-69	-24	4.9	1
Nonmatching – Matching: Sentence	F3op/F3t	44/45	R	54	18	30	5.2	14
	LPMC/F3op	6/44	L	-45	9	30	4.8	1
	ACC	6/8/32	M	6	27	42	6.9	52
Nonmatching – Matching: String	F3op/F3t	44/45	R	54	18	30	5.3	21
			R	39	18	33	4.7	1
	SMG	40	R	42	-30	48	5.0	2
Reverse"	LPMC	6	R	27	-9	51	4.7	1

doi:10.1371/journal.pone.0056230.t004

consistent with the signal changes in both Nested" and Reverse" (Table 2). The number of encoding might explain the results of Figure 4F, but its estimates cannot consistently explain the results of Figure 4D.

Effective and Anatomical Connectivity between L. F3op/F3t and L. SMG

Based on these results, we modeled effective connectivity between L. F3op/F3t and L. SMG by using DCM. Our interest was to identify the direction of the connectivity modulated by the Nested condition with largest DoM among all conditions, and the models were grouped into three "modulatory families": families A, B, and C, corresponding to the modulation for the bottom-up connection from L. SMG to L. F3op/F3t, for the top-down connection from L. F3op/F3t to L. SMG, and for both connections, respectively. Here we assumed intrinsic, i.e., task-independent, bidirectional connections. Each family was composed of three "input models" as regards the regions receiving driving inputs (see Figure S4 for all DCM models tested). Using a random-effects Bayesian model selection (BMS), we found that the family A was the most likely family (expected probability = 0.66, exceedance probability = 0.85) (Figure 4G and 4H). According to a second BMS for the input models within the family A, the model A1, in which L. F3op/F3t received driving inputs, was the best and highly probable model (expected probability = 0.77, exceedance probability = 0.95). For this particular model, we further tested whether the parameter estimates were significantly different from zero. The intrinsic connection from L. F3op/F3t to L. SMG was significantly positive [+0.22; one-sample *t*-test, *t*(17) = 4.8, *P* < 0.0002] (significance level at $\alpha = 0.025$, Bonferroni corrected within a parameter class of intrinsic connections) (Figure 4I), indicating that this top-down connection was consistent among the participants. The modulatory effect for the bottom-up connection was inhibitory [-0.17; *t*(17) = 1.4, *P* = 0.17], though it did not reach the significance level.

To further confirm the anatomical plausibility of the network between L. F3op/F3t and L. SMG revealed by DCM, we used DTI with a probabilistic tractography. Seed masks were set in the pair of L. F3op/F3t and L. SMG, both of which were significantly activated in Nested_(L)>Simple_(S). We identified a single contin-

uous cluster of the left SLF/AF that connected these regions (cluster size, 3,189 mm³), together with much smaller clusters or islands (Figure 4J). Moreover, the left SLF/AF was consistently observed in all participants.

Modulation of the Right Frontal Activations by Nonlinguistic Factors

We further examined the involvement of any error-related factors, which were residual factors that might induce cortical activation or deactivation. It should be noted that the factors listed in Tables 1 and 2 were equivalent between the matching and nonmatching stimuli. The [Nonmatching – Matching] contrast under either the sentence conditions (i.e., [Nested+Simple+Conjoined]) or the string conditions (i.e., [Reverse+Same]) consistently resulted in right-dominant activation, especially in R. F3op/F3t (Figure 5A and 5B), which was in accordance with the same demand of the matching task (Figure 1C and 1D). Other significantly activated regions were L. LPMC/F3op and ACC under the sentence conditions, as well as R. SMG under the string conditions (Table 4). As regards the [Matching – Nonmatching] contrast, no significant activation was seen under sentence or string conditions.

We also examined the activation in Reverse" for the effect of matching orders (e.g., A₂ A₁ B₁ B₂ vs. A₁ A₂ B₁ B₂; Figures 1B and 2B). The significant activation was observed only in R. LPMC (Figure 5C and Table 4), which suggested that activations could indeed be estimated by one and only non-null factor of memory span in Reverse" (Table 2). In Nested", the signal changes in R. LPMC were also significant (one-sample *t*-test, *P* < 0.05), which were not significantly different between Nested" and Reverse" (paired *t*-test, *P* = 0.98) (Figure 5D). This result was consistent with the equivalent estimates of memory span between Nested" and Reverse". It should be noted that R. LPMC activation was also observed for the main effects of condition and length (Figure 4A and 4B), which probably reflected the factor of memory span.

Discussion

By employing a novel paradigm to directly contrast jabberwocky sentences (Nested, Simple, and Conjoined) with letter

Table 5. Fittings and likelihood of various models tested.

L. F3op/F3t	Factor	RSS	r ²	P values for four contrasts	Log-likelihood	Likelihood ratio
	*DoM	0.0007	0.99	0.17, 0.92, 0.97, 0.99	65.0	1.0
	DoM+No. of Search	0.0065	0.88	0.0035, 0.064, 0.63, 0.88	59.2	3.1 × 10 ⁻³
	No. of Search	0.052	<0.1	<0.0001, 0.018, 0.019, 0.031	33.4	2.0 × 10 ⁻¹⁴
	No. of Merge	0.053	0	<0.0001, 0.0035, 0.018, 0.17	n/a	n/a
	No. of case markers (-ga/-no)	0.053	0	<0.0001, 0.0035, 0.018, 0.17	n/a	n/a
	No. of tense markers (-ru/-ta)	0.0067	0.87	0.0035, 0.17, 0.32, 0.56	59.7	4.8 × 10 ⁻³
	Degree of nesting	0.010	0.80	0.0035, 0.018, 0.17, >0.99	57.1	3.7 × 10 ⁻⁴
	Degree of self-embedding	0.015	0.71	0.0035, 0.0075, 0.019, 0.17	53.3	8.7 × 10 ⁻⁶
	No. of nodes	0.015	0.72	0.0050, 0.0082, 0.018, 0.17	53.7	1.2 × 10 ⁻⁵
	Depth of postponed symbols	0.053	0	<0.0001, 0.0035, 0.018, 0.17	n/a	n/a
	Integration costs	0.0066	0.88	0.0017, 0.15, 0.48, 0.53	59.0	2.5 × 10 ⁻³
	Storage costs	0.014	0.74	<0.0001, 0.024, 0.83, 0.85	53.8	1.3 × 10 ⁻⁵
	Syntactic interference	0.0067	0.87	0.0035, 0.17, 0.32, 0.56	59.7	4.8 × 10 ⁻³
	Positional similarity	0.0055	0.90	0.051, 0.12, 0.17, 0.19	60.1	7.8 × 10 ⁻³
	Memory span	0.0066	0.88	0.0017, 0.15, 0.48, 0.53	59.0	2.5 × 10 ⁻³
	Counting	0.017	0.67	0.0003, 0.0013, 0.035, 0.72	50.8	7.0 × 10 ⁻⁷
	No. of encoding	0.051	<0.1	<0.0001, 0.014, 0.018, 0.12	32.9	1.2 × 10 ⁻¹⁴
	Memory span+counting	0.0099	0.81	0.0007, 0.035, 0.15, 0.76	55.5	7.9 × 10 ⁻⁵
	Memory span+No. of encoding	0.015	0.72	<0.0001, 0.10, 0.46, 0.59	52.5	3.6 × 10 ⁻⁶
L. SMG	Factor	RSS	r ²	P values for four contrasts	Log-likelihood	Likelihood ratio
	DoM	0.0063	0.92	0.013, 0.083, 0.44, 0.49	58.8	0.079
	*DoM+No. of Search	0.0020	0.97	0.22, 0.30, 0.42, 0.62	61.4	1.0
	No. of Search	0.075	<0.1	<0.0001, 0.0061, 0.045, 0.090	23.6	3.8 × 10 ⁻¹⁷
	No. of Merge	0.076	0	<0.0001, 0.0061, 0.013, 0.22	n/a	n/a
	No. of case markers (-ga/-no)	0.076	0	<0.0001, 0.0061, 0.013, 0.22	n/a	n/a
	No. of tense markers (-ru/-ta)	0.0079	0.90	0.013, 0.023, 0.22, 0.34	55.9	4.1 × 10 ⁻³
	Degree of nesting	0.0088	0.88	0.0061, 0.013, 0.22, >0.99	55.5	2.8 × 10 ⁻³
	Degree of self-embedding	0.023	0.69	0.0002, 0.0018, 0.013, 0.22	45.5	1.2 × 10 ⁻⁷
	No. of nodes	0.033	0.56	0.0004, 0.0005, 0.0061, 0.013	40.1	6.0 × 10 ⁻¹⁰
	Depth of postponed symbols	0.076	0	<0.0001, 0.0061, 0.013, 0.22	n/a	n/a
	Integration costs	0.021	0.72	0.0001, 0.014, 0.028, 0.18	46.3	2.7 × 10 ⁻⁷
	Storage costs	0.032	0.58	<0.0001, 0.0014, 0.084, 0.49	40.3	7.1 × 10 ⁻¹⁰
	Syntactic interference	0.0079	0.90	0.013, 0.023, 0.22, 0.34	55.9	4.1 × 10 ⁻³
	Positional similarity	0.020	0.73	0.0039, 0.0052, 0.013, 0.029	47.6	1.0 × 10 ⁻⁶
	Memory span	0.021	0.72	0.0001, 0.014, 0.028, 0.18	46.3	2.7 × 10 ⁻⁷
	Counting	0.041	0.46	<0.0001, <0.0001, 0.0039, 0.77	35.6	6.2 × 10 ⁻¹²
	No. of encoding	0.076	<0.1	<0.0001, 0.0061, 0.017, 0.16	22.5	1.4 × 10 ⁻¹⁷
	Memory span+counting	0.028	0.63	<0.0001, 0.0018, 0.0086, 0.44	41.3	1.9 × 10 ⁻⁹
	Memory span+No. of encoding	0.011	0.85	0.0034, 0.051, 0.13, 0.81	52.1	9.7 × 10 ⁻⁵

Percent signal changes in L. F3op/F3t and L. SMG were fitted with a single scale parameter to a model of each factor using its subtracted estimates (Table 1) for the four contrasts of Nested_(L), Nested_(S), Simple_(L), and Simple_(S). The P values for the t-tests are shown in ascending order. Note that the models of DoM and “DoM+number of Search” (with an asterisk) resulted in the best fit for L. F3op/F3t and L. SMG, respectively, i.e., with the least residual sum of squares (RSS), largest coefficient of determination (r²), and larger P values. The likelihood of models with all null estimates was incalculable (n/a). A likelihood ratio is the ratio of each model’s likelihood to the best model’s likelihood. The best models of DoM and “DoM+number of Search” for L. F3op/F3t and L. SMG, respectively, were by far more likely than the other models.

doi:10.1371/journal.pone.0056230.t005

strings (Reverse and Same) (Figures 1 and 2), we obtained four striking results. First, we found that DoM was indeed a key syntactic factor that could account for syntax-selective activations in L. F3op/F3t and L. SMG, localized by the Nested_(L)>Sim-

ple_(S) contrast (Figure 4C and 4D). By constructing a model of each syntactic, other linguistic, or nonlinguistic factor using its estimates (Table 1), we demonstrated that the models of DoM and “DoM+number of Search” were the best to explain L. F3op/F3t

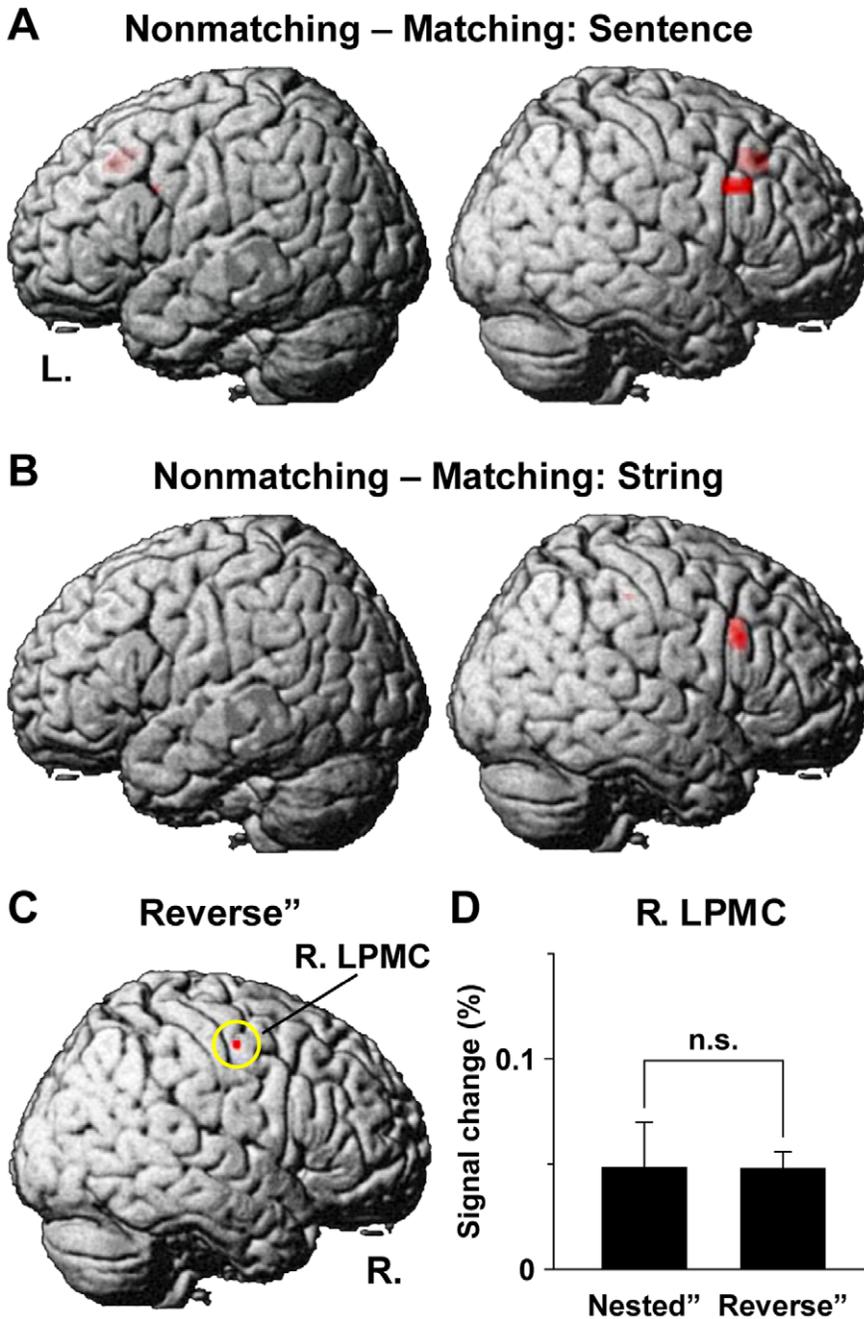


Figure 5. Modulation of the right frontal activations by nonlinguistic factors. One-sample *t*-tests were used for the contrasts indicated. (A) Regions identified by [Nonmatching – Matching] under the sentence conditions, related to error-related factors. Note the right-dominant activation, especially in R. F3op/F3t. (B) Regions identified by [Nonmatching – Matching] under the string conditions. (C) Regions identified by Reverse”. This contrast revealed the difference in matching orders (e.g., A₂ A₁ B₁ B₂ vs. A₁ A₂ B₁ B₂). Note the significant activation in R. LPMC. (D) The percent signal changes in R. LPMC, which was consistent with the equivalent estimates of memory span (see Table 2). doi:10.1371/journal.pone.0056230.g005

and L. SMG activations, respectively (Table 5). Secondly, by directly contrasting jabberwocky sentences with letter strings, i.e., Nested”>Reverse”, we showed that the selective activation in L. F3op/F3t and L. SMG, which was consistent with the involvement of the syntactic factors demonstrated above, was replicated irrespective of identical matching orders and symbol orders (e.g., N₂ N₁ V₁ V₂ and A₂ A₁ B₁ B₂ for the Nested and Reverse, respectively) (Figure 4E and 4F). This point is particularly important, because temporal order-related or memory-related

factors have often been confused with differences in structure or grammar type. Our results strongly support that syntactic structures are recursively constructed when well-formed sentences are given. Thirdly, by using DCM, we found that the model with an inhibitory modulatory effect for the bottom-up connectivity from L. SMG to L. F3op/F3t, and with driving inputs to L. F3op/F3t, was the best one (Figure 4G and 4H). For this best model, the top-down connection from L. F3op/F3t to L. SMG was significantly positive (Figure 4I). By using DTI, we also confirmed that the left

dorsal pathway of SLF/AF consistently connected these two regions (Figure 4J). These results suggest that there is a transmission of information about DoM through this specific dorsal pathway. Lastly, we established that nonlinguistic order-related and error-related factors significantly activated mostly right frontal regions. The difference in memory span significantly modulated R. LPMC activation in Reverse”, suggesting that this region plays a major role in tracking matching orders (Figure 5C and 5D), while error-related factors in [Nonmatching – Matching] consistently modulated R. F3op/F3t activation under both sentence and string conditions (Figure 5A and 5B). In summary, these results indicate that the identified network of L. F3op/F3t and L. SMG subserves the calculation of DoM in recursively merged sentences, and that R. LPMC monitors memory span to drive a memory-maintenance system. If multiple factors, such as the number of nodes, memory span, etc., are equally plausible to explain activations, then a superordinate concept, such as “syntactic complexity”, can be a more useful factor than individual factors. However, in the present experiment, the minimal factor of DoM *sufficiently* explained the activation pattern observed, while other factors were by far less likely (see Table 5). Therefore, syntactic complexity was restricted and replaced by DoM as a more fundamental concept, just like the historical development from “gene” to DNA.

It remains a central issue in cognitive sciences whether or not the faculty of language is also shared by animals. Animals have been thus tested with regular symbolic sequences such as $A^n B^n$ ($n \geq 2$; i.e., AABB, AAABBB, ...) and $(AB)^n$ ($n \geq 2$; i.e., ABAB, ABABAB, ...), which differ in *symbol order*. In an animal study, songbirds were trained to discriminate patterns of $A^n B^n$ and $(AB)^n$ in more than ten thousand trials [32]. However, this learning can be achieved by a counting strategy alone [33]. There is also a recent report that songbirds seemed to discriminate strings with or without nesting [34], but this learning can be achieved by simply remembering partial strings [35]. Along the line of contrasting $A^n B^n$ and $(AB)^n$, fMRI studies have tested participants with different symbolic sequences, such as $A_2 A_1 B_1 B_2$ versus $A_1 B_1 A_2 B_2$, which also differ in *matching order* [36]. However, the difference in activation patterns can be simply explained by differences in any factors associated with matching orders and symbol orders, i.e., temporal order-related factors. It was thus necessary to completely control these general factors when extracting any syntactic factors from a number of cognitive factors involved in actual symbol processing.

Our finding that L. F3op/F3t subserves the syntactic computation further extends the functional specialization of this region reported previously [14,18–20]. Some previous fMRI studies have interpreted L. F3op/F3t activation as reflecting temporal order-related or memory-related factors [37,38]. However, these previous studies contrasted hierarchically complex sentences with simpler sentences, while it is clear that syntactic factors, including DoM, were also involved. Moreover, the previously reported modulation of the L. F3op/F3t activation by scrambling word orders [3] can be consistently explained by DoM, because scrambling requires “movements” of NPs to higher nodes by applying more Merge operations, thus increasing DoM. The size of linguistic constituents also correlates with DoM, especially when the number of left/right branches was increased as in the case of Pallier et al. (2011) [4]. In the present study, we characterized the neural substrates of syntactic computation by segregating a number of possible factors, and demonstrated that the exact activations in L. F3op/F3t can be used to calculate DoM. Indeed, each structure of our jabberwocky sentences was uniquely represented by DoM, together with the numbers of Merge and Search (see Table 1).

A previous fMRI study involving the implicit learning of an artificial regular grammar has reported that the “ungrammatical – grammatical” contrast for symbol sequences activated L. F3op/F3t, suggesting that such activation was due to artificial syntactic violations among any error-related factors [39]. However, this result may not depend on the presence of errors themselves, but on other rule-related processes associated with error-correction, etc. In contrast, we have previously demonstrated that an explicit syntactic decision enhanced L. F3op/F3t activation under *both* grammatical and ungrammatical conditions [17]. On the other hand, a recent fMRI study has compared nested and branching constructions, suggesting that activation in the bilateral posterior superior temporal cortex reflects an integration of lexico-semantic and syntactic information [40]. However, as regards this previous result, the effects of semantic factors were inevitably confounded with any structural processing, because real German sentences were used as stimuli in that study. Furthermore, according to our paradigm, the temporal cortex in neither of the hemispheres showed any significant activation for the Nested (Figure 4). It was thus quite important to verify that activation in L. F3op/F3t, but not in the temporal cortex, is indeed crucial for syntactic processing.

In the present study, we found that L. SMG activations were modulated by “DoM+number of Search”. Consistent with the suggested role of L. AG/SMG for vocabulary knowledge or lexical processing [21,22], the number of Search is likely to induce such a modulation, in the sense that Search assigns a specific feature that can be linked with morphosyntactic changes. The Japanese language happens to lack the agreement of grammatical features, but it is nevertheless equipped with the general Search procedure attested for various phenomena in the language [8]. Our results suggest that Search actually applied to a subject-verb pair of a jabberwocky sentence in the present paradigm, where the relevant features (vowels here) are experimentally “inserted”. It should also be noted in this connection that the Japanese language exhibits a phenomenon called “honorification” (the case of an honored person and the form of honorifics on verbs optionally match) [41,42], in which Search assigns such features as honorifics. Our previous fMRI study using an honorification judgment task reported activation in L. F3op/F3t and L. LPMC, as well as in the L. inferior parietal gyrus and L. AG [43], which is consistent with activation in L. AG/SMG in the present study (Tables 3 and 4).

Our DCM and DTI results further indicate that L. SMG activations reflecting DoM mirrored a top-down influence from L. F3op/F3t through the left dorsal pathway of SLF/AF. A recent DCM study with a picture-sentence matching task has suggested that L. F3op/F3t received driving inputs [44], which was consistent with our DCM results. Moreover, our previous studies revealed that the functional connectivity between L. F3t/F3O (pars orbitalis) and L. AG/SMG was selectively enhanced during sentence processing [45], and that L. AG/SMG was also activated during the identification of correct past-tense forms of verbs, probably reflecting an integration of syntactic and vocabulary knowledge [46]. Considering the role of L. AG/SMG in lexical processing, the Search operation based on DoM would be essential in assigning relevant features to the syntactic objects derived from lexical items.

In [Nonmatching – Matching], R. F3op/F3t was consistently activated under both sentence and string conditions (Figure 5A and 5B), whereas L. LPMC/F3op, ACC, or R. SMG were activated under either condition. These four regions were also activated in Nested’>Simple’, and in Long>Short while combining Nested’ and Simple’; the ACC and R. SMG were activated in Nested’>Reverse’ as well. It appears likely that a part of the activation in these four regions reflects error-related factors including the detection and correction of errors, which would be more demanding with the

Nested, as well as in the Long>Short contrast. Because L. LPMC has been known to selectively subserve syntactic processing [15,19,47], a weak activation in L. LPMC/F3op only under the sentence conditions may reflect the confirmation of sentence constructions when confronted with nonmatching stimuli. On the other hand, it has been suggested that the dorsal ACC plays a major role during conflict monitoring during a highly demanding task, e.g., a Stroop task [48]. Our recent magnetoencephalography study also suggested that the anterior portion of the ACC is a candidate region for monitoring syntactically anomalous sentences [49]. Moreover, previous studies on a response inhibition, typically tested with a Nogo task, suggested that R. F3op/F3t, ACC, and R. SMG were also involved in monitoring anomalous stimuli [50]. In contrast to these factors that activated mostly right and medial regions, it is noteworthy that the syntactic factors clearly activated the left frontal and parietal regions.

Any factors associated with matching orders and symbol orders might influence activation in the language areas, but we clearly showed that R. LPMC was activated in Reverse" (Figure 5C) for the effect of memory span related to matching orders. The study of real German sentences also reported activation in the right dorsal premotor area for the contrast of nested vs. branching constructions [40], but the right dorsal premotor area was not the same region as R. LPMC in the present study. In this German study, memory span was controlled by the insertion of some words, while matching orders and symbol orders still differed, and thus factors other than memory span were inevitably introduced to interpret the right dorsal premotor activation. The identification of critical factors in language processing thus inevitably depends on an experimental design that involves an effective contrast of conditions. One promising direction of research is to further clarify activations modulated by other linguistic and nonlinguistic factors, which may eventually make possible the elucidation of all aspects of linguistic information in the human brain.

Supporting Information

Figure S1 Application of other structure-based models to sentences with complex structures, I. (A) The digits shown in red and blue denote "degree of nesting" and "degree of self-embedding", respectively. Nested and self-embedded constructions occur within sentences (Ss). Note that each shortest "zigzag path" counts one for the degree of nesting or self-embedding. For the Nested_(L), S₁ dominates [N₂ S₂ V₂], and S₀ in turn dominates [N₃ S₁ V₃], i.e., [N₃[N₂ S₂ V₂]V₃]; the degree of nesting or self-embedding is thus two (the number of blue dots minus one). For the Simple_(L), both of (NN)₁N₁ and N(NN)₁ yield the same maximum degree of nesting or self-embedding for an entire sentence. (B) The digits shown in red denote "number of nodes". (TIF)

Figure S2 Application of other structure-based models to sentences with complex structures, II. The digits shown in red and blue denote the number of branches from each node and "depth of postponed symbols" [51], respectively. The largest estimate can be obtained by adding together the digits shown in red with circles. For the Simple_(L), the largest estimate of "depth of postponed symbols" is obtained, when Vs take a right-branching construction of V₁(VV₁). For the Conjoined_(L), the depth of postponed symbols is increased by two to reach the rightmost branches, when conjoining three sentences at a multiple-branching node. (TIF)

Figure S3 Application of other structure-based models to sentences with complex structures, III. (A) The digits shown in red and blue denote "integration costs" and "storage costs" [52], respectively. Integration costs are estimated at every stimulus by adding together "new discourse referents" and "structural integrations". For example, at V₂ of the Nested_(L), N₁ and V₁ intervene while making [N₂[N₁ V₁]V₂] (structural integrations = 2), and one verb completes the input with *-to* or *-te* (storage cost = 1). Note that the estimate of maximum structural integrations in a sentence matches with that of memory span in our paradigm. (B) The digits shown in red and blue denote "syntactic interference" and "positional similarity" [53], respectively. Syntactic interference is estimated at every stimulus by adding together "retroactive interference" and "proactive interference". For example, at V₂ of the Nested_(L), the attachment of V₂ to N₂ suffers from one unit of retroactive interference from N₁, and from one unit of proactive interference from N₃ (syntactic interference = 2). There are three adjacent nominative NPs in this sentence (positional similarity = 3). (TIF)

Figure S4 The DCM models tested. We assumed bi-directional connectivity between L. F3op/F3t and L. SMG. The models were grouped into three modulatory families based on the modulations of the connections under the Nested condition: Family A (A1–A3), in which the connection from L. SMG to L. F3op/F3t was modulated, Family B (B1–B3), in which the connection from L. F3op/F3t to L. SMG was modulated, and Family C (C1–C3), in which both connections were modulated. Each family was composed of three "input models" as regards the regions receiving driving inputs. (TIF)

Table S1 Examples of short nonmatching stimuli. (PDF)

Table S2 Examples of long nonmatching stimuli. (PDF)

Appendix S1 Theoretical issues. (PDF)

Appendix S2 Detailed information about the stimuli. (PDF)

Appendix S3 Task instructions and training procedures. (PDF)

Appendix S4 MRI data analyses. (PDF)

Acknowledgments

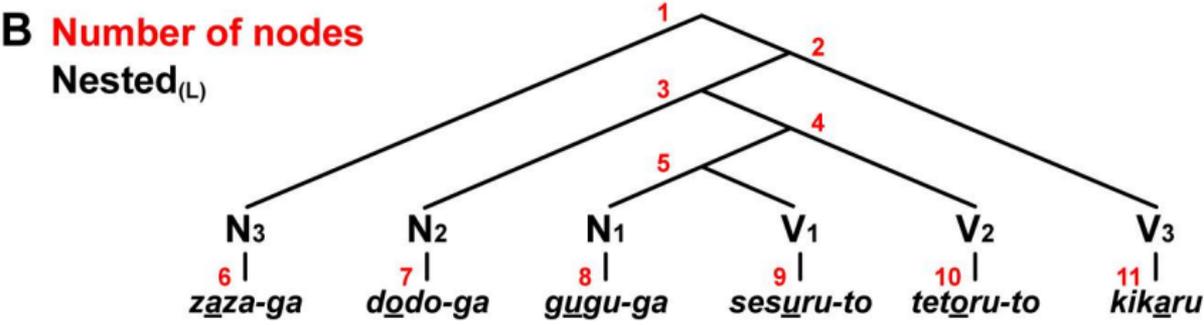
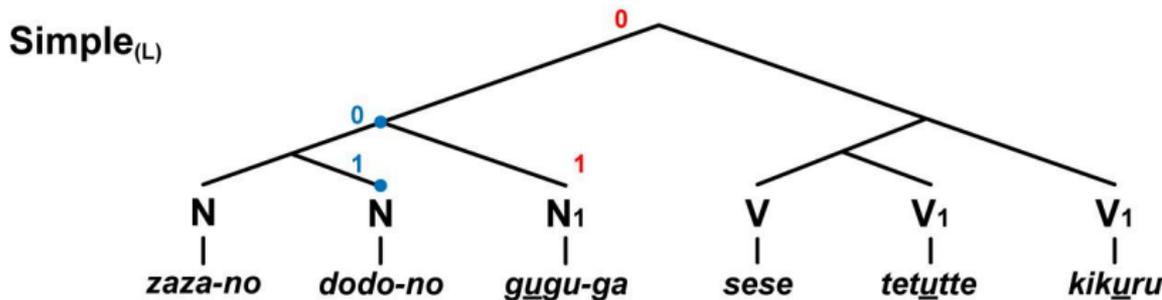
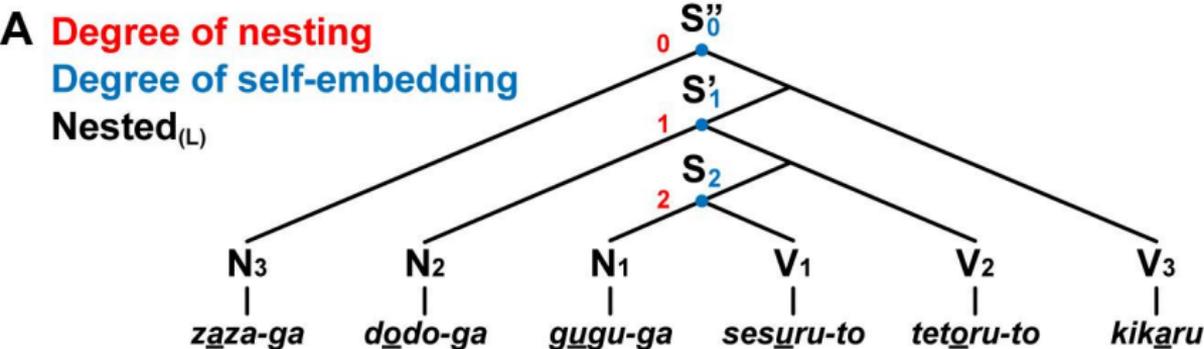
This paper is dedicated to the memory of S.-Y. Kuroda, who was a truly distinguished linguist, and a great mentor for all of us in the field. We would like to thank A. Sakurai and A. Morita for their help in earlier experiments, H. Miyashita, K. Iijima, and T. Inubushi for their comments on the earlier manuscript, N. Komoro and N. Saeki for their technical assistance, and H. Matsuda and S. Matsukura for their administrative assistance.

Author Contributions

Conceived and designed the experiments: SO NF KLS. Performed the experiments: SO NF KLS. Analyzed the data: SO KLS. Wrote the paper: SO NF KLS.

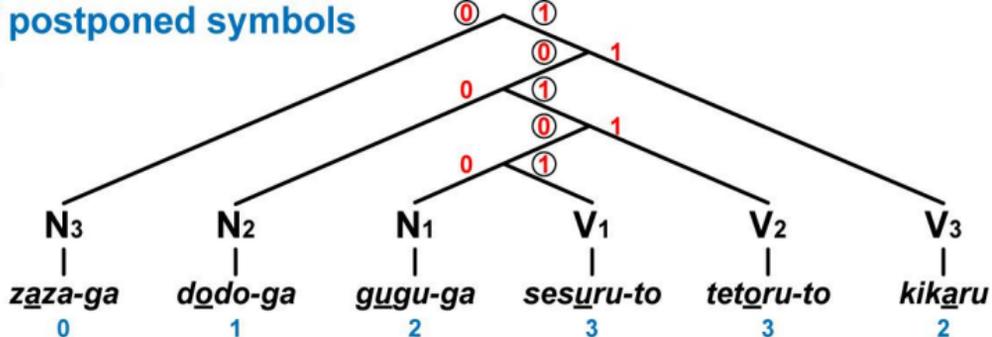
References

1. Hopcroft JE, Ullman JD (1979) Introduction to Automata Theory, Languages, and Computation. Reading, MA: Addison-Wesley. 418 p.
2. O'Grady W, Archibald J, Aronoff M, Rees-Miller J (2010) Contemporary Linguistics: An Introduction, 6th Edition. Boston, MA: Bedford/St. Martin's. 694 p.
3. Röder B, Stock O, Neville H, Bien S, Röslér F (2002) Brain activation modulated by the comprehension of normal and pseudo-word sentences of different processing demands: A functional magnetic resonance imaging study. *Neuroimage* 15: 1003–1014.
4. Pallier C, Devauchelle A-D, Dehaene S (2011) Cortical representation of the constituent structure of sentences. *Proc Natl Acad Sci USA* 108: 2522–2527.
5. Chomsky N (1957) Syntactic Structures. The Hague: Mouton Publishers. 117 p.
6. Chomsky N (1965) Aspects of the Theory of Syntax. Cambridge, MA: The MIT Press. 251 p.
7. Chomsky N (1995) The Minimalist Program. Cambridge, MA: The MIT Press. 420 p.
8. Fukui N, Sakai H (2003) The visibility guideline for functional categories: Verb raising in Japanese and related issues. *Lingua* 113: 321–375.
9. Fukui N (2011) Merge and Bare Phrase Structure. In: Boeckx C, editor. The Oxford Handbook of Linguistic Minimalism. Oxford, UK: Oxford University Press. 73–95.
10. Saito M, Fukui N (1998) Order in phrase structure and movement. *Ling Inq* 29: 439–474.
11. Chomsky N (2004) Beyond Explanatory Adequacy. In: Belletti A, editor. Structures and Beyond: The Cartography of Syntactic Structures, Volume 3. Oxford, UK: Oxford University Press. 104–131.
12. Stromswold K, Caplan D, Alpert N, Rauch S (1996) Localization of syntactic comprehension by positron emission tomography. *Brain Lang* 52: 452–473.
13. Dapretto M, Bookheimer SY (1999) Form and content: Dissociating syntax and semantics in sentence comprehension. *Neuron* 24: 427–432.
14. Embick D, Marantz A, Miyashita Y, O'Neil W, Sakai KL (2000) A syntactic specialization for Broca's area. *Proc Natl Acad Sci USA* 97: 6150–6154.
15. Hashimoto R, Sakai KL (2002) Specialization in the left prefrontal cortex for sentence comprehension. *Neuron* 35: 589–597.
16. Friederici AD, Rüschemeyer S-A, Hahne A, Fiebach CJ (2003) The role of left inferior frontal and superior temporal cortex in sentence comprehension: Localizing syntactic and semantic processes. *Cereb Cortex* 13: 170–177.
17. Suzuki K, Sakai KL (2003) An event-related fMRI study of explicit syntactic processing of normal/anomalous sentences in contrast to implicit syntactic processing. *Cereb Cortex* 13: 517–526.
18. Musso M, Moro A, Glauche V, Rijntjes M, Reichenbach J, et al. (2003) Broca's area and the language instinct. *Nat Neurosci* 6: 774–781.
19. Kinno R, Kawamura M, Shioda S, Sakai KL (2008) Neural correlates of noncanonical syntactic processing revealed by a picture-sentence matching task. *Hum Brain Mapp* 29: 1015–1027.
20. Sakai KL (2005) Language acquisition and brain development. *Science* 310: 815–819.
21. Lee H, Devlin JT, Shakeshaft C, Stewart LH, Brennan A, et al. (2007) Anatomical traces of vocabulary acquisition in the adolescent brain. *J Neurosci* 27: 1184–1189.
22. Pattamadilok C, Knierim IN, Duncan KJK, Devlin JT (2010) How does learning to read affect speech perception? *J Neurosci* 30: 8435–8444.
23. Friston KJ, Henson RN (2006) Commentary on: Divide and conquer; A defence of functional localisers. *Neuroimage* 30: 1097–1099.
24. Friston KJ, Harrison L, Penny W (2003) Dynamic causal modelling. *Neuroimage* 19: 1273–1302.
25. Saur D, Kreher BW, Schnell S, Kümmerer D, Kellmeyer P, et al. (2008) Ventral and dorsal pathways for language. *Proc Natl Acad Sci USA* 105: 18035–18040.
26. Wong FCK, Chandrasekaran B, Garibaldi K, Wong PCM (2011) White matter anisotropy in the ventral language pathway predicts sound-to-word learning success. *J Neurosci* 31: 8780–8785.
27. Wilson SM, Galantucci S, Tartaglia MC, Rising K, Patterson DK, et al. (2011) Syntactic processing depends on dorsal language tracts. *Neuron* 72: 397–403.
28. Griffiths JD, Marslen-Wilson WD, Stamatakis EA, Tyler LK (2013) Functional organization of the neural language system: Dorsal and ventral pathways are critical for syntax. *Cereb Cortex* 23: 139–147.
29. Oldfield RC (1971) The assessment and analysis of handedness: The Edinburgh inventory. *Neuropsychologia* 9: 97–113.
30. Tsujimura N (2007) An Introduction to Japanese Linguistics, 2nd Edition. Malden, MA: Blackwell Publishing. 501 p.
31. Kriegeskorte N, Simmons WK, Bellgowan PSF, Baker CI (2009) Circular analysis in systems neuroscience: The dangers of double dipping. *Nat Neurosci* 12: 535–540.
32. Gentner TQ, Fenn KM, Margoliash D, Nusbaum HC (2006) Recursive syntactic pattern learning by songbirds. *Nature* 440: 1204–1207.
33. Corballis MC (2007) Recursion, language, and starlings. *Cogn Sci* 31: 697–704.
34. Abe K, Watanabe D (2011) Songbirds possess the spontaneous ability to discriminate syntactic rules. *Nat Neurosci* 14: 1067–1074.
35. Beckers GJL, Bolhuis JJ, Okanoya K, Berwick RC (2012) Birdsong neurolinguistics: Songbird context-free grammar claim is premature. *NeuroReport* 23: 139–145.
36. Bahlmann J, Schubotz RI, Friederici AD (2008) Hierarchical artificial grammar processing engages Broca's area. *Neuroimage* 42: 525–534.
37. Bahlmann J, Rodriguez-Fornells A, Rotte M, Münte TF (2007) An fMRI study of canonical and noncanonical word order in German. *Hum Brain Mapp* 28: 940–949.
38. Santi A, Grodzinsky Y (2010) fMRI adaptation dissociates syntactic complexity dimensions. *Neuroimage* 51: 1285–1293.
39. Petersson KM, Forkstam C, Ingvar M (2004) Artificial syntactic violations activate Broca's region. *Cogn Sci* 28: 383–407.
40. Friederici AD, Makuuchi M, Bahlmann J (2009) The role of the posterior superior temporal cortex in sentence comprehension. *NeuroReport* 20: 563–568.
41. Gunji T (1987) Japanese Phrase Structure Grammar: A Unification-Based Approach. Dordrecht: D. Reidel Publishing Company. 239 p.
42. Ivana A, Sakai H (2007) Honorification and light verbs in Japanese. *J East Asian Ling* 16: 171–191.
43. Momo K, Sakai H, Sakai KL (2008) Syntax in a native language still continues to develop in adults: Honorification judgment in Japanese. *Brain Lang* 107: 81–89.
44. den Ouden D-B, Saur D, Mader W, Schelter B, Lukic S, et al. (2012) Network modulation during complex syntactic processing. *Neuroimage* 59: 815–823.
45. Homae F, Yahata N, Sakai KL (2003) Selective enhancement of functional connectivity in the left prefrontal cortex during sentence processing. *Neuroimage* 20: 578–586.
46. Tatsuno Y, Sakai KL (2005) Language-related activations in the left prefrontal regions are differentially modulated by age, proficiency, and task demands. *J Neurosci* 25: 1637–1644.
47. Indefrey P, Hagoort P, Herzog H, Seitz RJ, Brown CM (2001) Syntactic processing in left prefrontal cortex is independent of lexical meaning. *Neuroimage* 14: 546–555.
48. Botvinick MM, Cohen JD, Carter CS (2004) Conflict monitoring and anterior cingulate cortex: An update. *Trends Cogn Sci* 8: 539–546.
49. Iijima K, Fukui N, Sakai KL (2009) The cortical dynamics in building syntactic structures of sentences: An MEG study in a minimal-pair paradigm. *Neuroimage* 44: 1387–1396.
50. Chikazoe J, Konishi S, Asari T, Jimura K, Miyashita Y (2007) Activation of right inferior frontal gyrus during response inhibition across response modalities. *J Cogn Neurosci* 19: 69–80.
51. Yngve VH (1960) A model and an hypothesis for language structure. *Proc Am Philos Soc* 104: 444–466.
52. Gibson E (2000) The Dependency Locality Theory: A Distance-Based Theory of Linguistic Complexity. In: Marantz A, Miyashita Y, O'Neil W, editors. Image, Language, Brain: Papers from the First Mind Articulation Project Symposium. Cambridge, MA: The MIT Press. 95–126.
53. Lewis R, Nakayama M (2002) Syntactic and positional similarity effects in the processing of Japanese embeddings. In: Nakayama M, editor. Sentence Processing in East Asian Languages. Stanford, CA: CSLI Publications. 85–111.

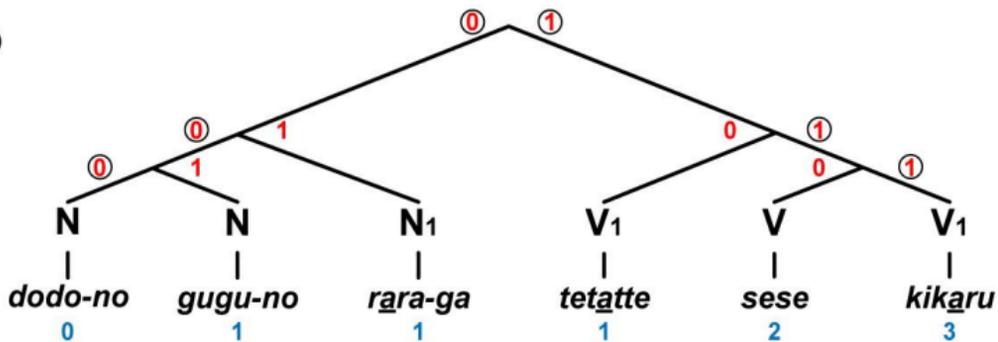


Depth of postponed symbols

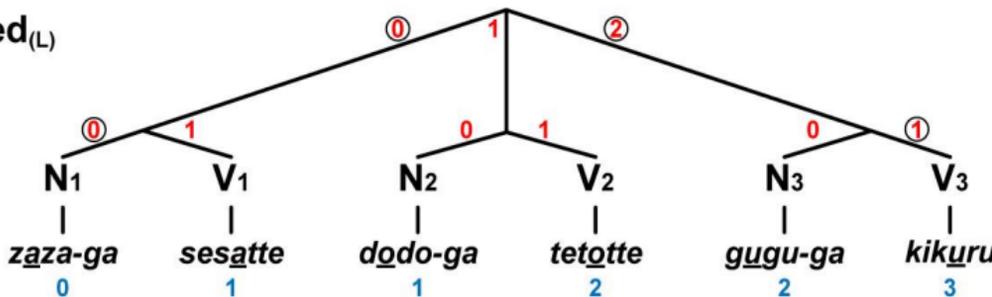
Nested_(L)



Simple_(L)

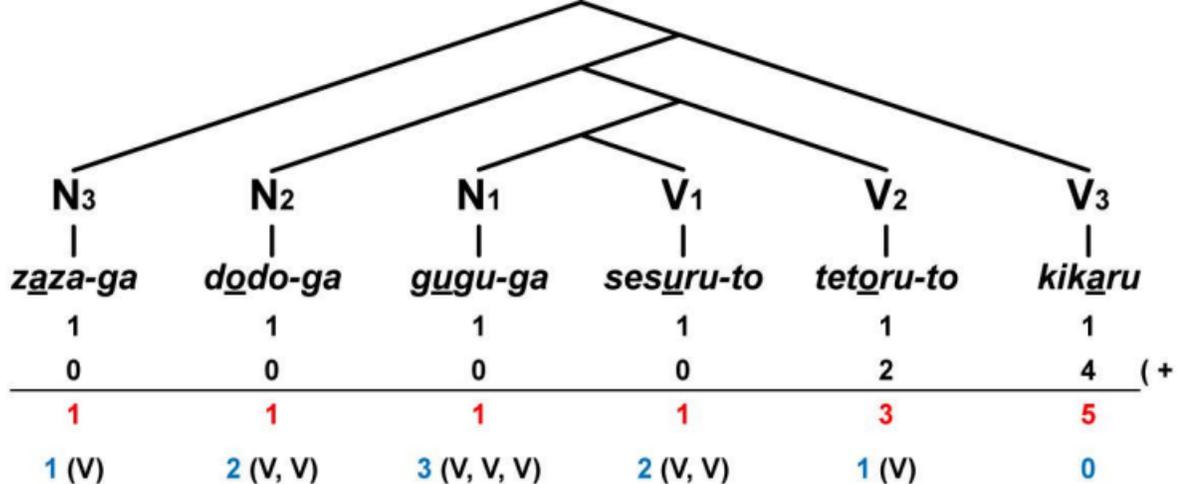


Conjoined_(L)



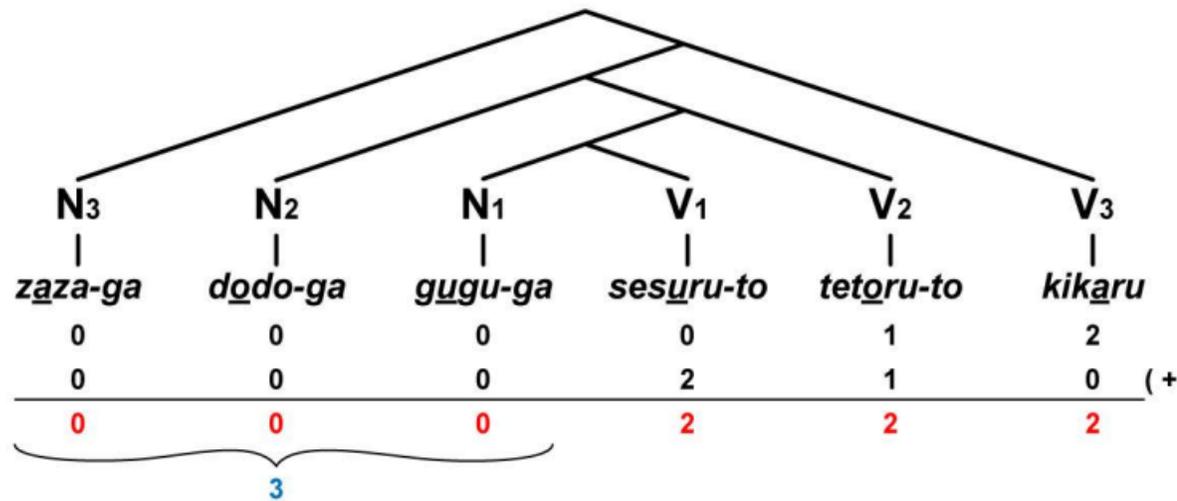
A Nested_(L)

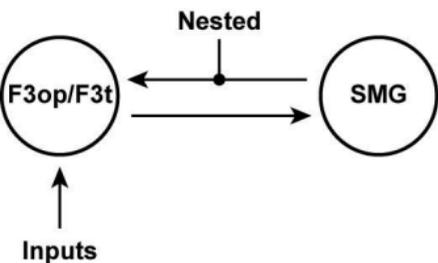
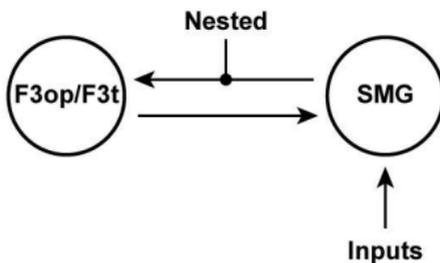
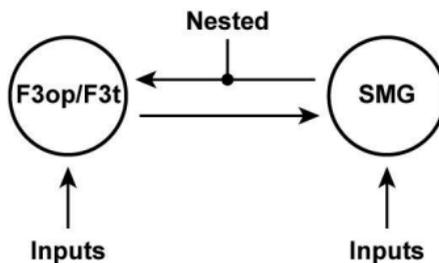
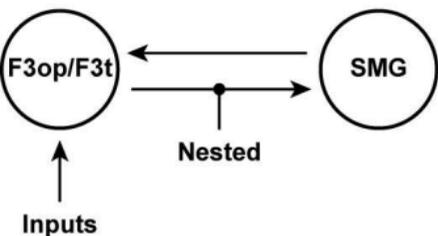
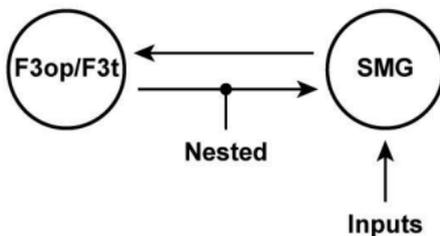
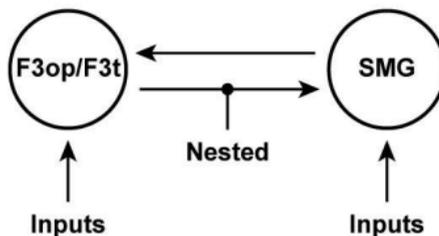
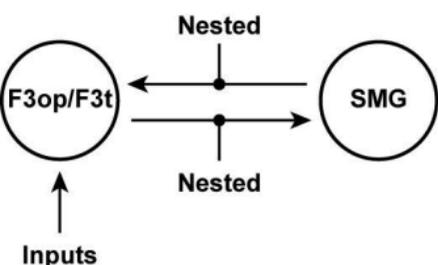
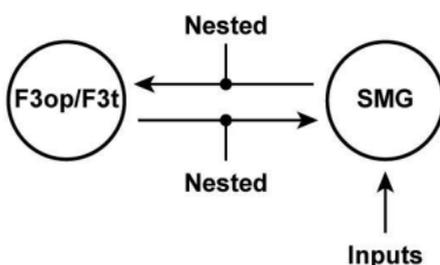
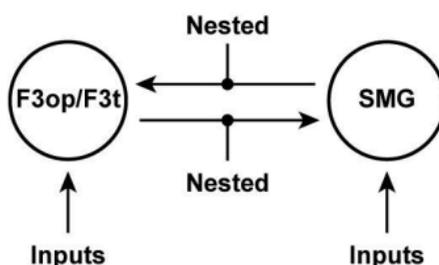
New discourse referents
 Structural integrations
Integration costs
Storage costs



B Nested_(L)

Retroactive interference
 Proactive interference
Syntactic interference
Positional similarity



Model A1**Model A2****Model A3****Model B1****Model B2****Model B3****Model C1****Model C2****Model C3**

Syntactic Computation in the Human Brain: The Degree of Merger as a Key Factor

Shinri Ohta, Naoki Fukui, Kuniyoshi L. Sakai

Table S1. Examples of short nonmatching stimuli.

Condition	Error type	Stimulus example
Nested sentence, short (Nested _(S))	$[N_2[N_1 V_1]V_2^*]$	<i>mumu-ga rara-ga tetaru-to hih<u>o</u>ru</i>
	$[N_2[N_1 V_1^*]V_2^*]$	<i>dodo-ga gugu-ga tet<u>o</u>ru-to kik<u>a</u>ru</i>
	$[N_2[N_1 V_1^*]V_2]$	<i>rara-ga mumu-ga ses<u>o</u>tta-to kikatta</i>
Simple sentence, short (Simple _(S))	$[(NN_1) (VV_1^*)]$	<i>rara-no gugu-ga tete ses<u>a</u>ttta</i>
Conjoined sentence, short (Conjoined _(S))	$[N_1 V_1][N_2 V_2^*]$	<i>rara-ga hihatte gugu-ga tet<u>o</u>ru</i>
	$[N_1 V_1^*][N_2 V_2^*]$	<i>yoyo-ga hih<u>u</u>tte rara-ga ses<u>o</u>ttta</i>
	$[N_1 V_1^*][N_2 V_2]$	<i>gugu-ga tet<u>o</u>tte zaza-ga hiharu</i>
Reverse-order string, short (Reverse _(S))	$A_2 A_1 B_1 B_2^*$	<i>nododo rukiku rukiku <u>donodo</u></i>
	$A_2 A_1 B_1^* B_2^*$	<i>gayoyo settasa <u>sassetta</u> <u>gadodo</u></i>
	$A_2 A_1 B_1^* B_2$	<i>mugamu sessota <u>kittako</u> mugamu</i>
Same-order string, short (Same _(S))	$A_1 A_2 B_1 B_2^*$	<i>ruteta gugagu ruteta <u>yogayo</u></i>
	$A_1 A_2 B_1^* B_2^*$	<i>yonoyo rusesu <u>donodo</u> <u>rususe</u></i>
	$A_1 A_2 B_1^* B_2$	<i>norara kuruki <u>noyoyo</u> kuruki</i>

For each condition, nonmatching stimuli with errors in different positions are listed in descending order of frequency. Under the sentence conditions, Vs with asterisks represent matching errors (underlined vowels in stimulus examples here; no underline in the real stimuli). Under the string conditions, Bs with asterisks represent matching errors (underlined letter strings in stimulus examples here; no underline in the real stimuli).

Syntactic Computation in the Human Brain:

The Degree of Merger as a Key Factor

Shinri Ohta, Naoki Fukui, Kuniyoshi L. Sakai

Table S2. Examples of long nonmatching stimuli.

Condition	Error type	Stimulus example
Nested _(L)	$[N_3[N_2[N_1 V_1]V_2]V_3^*]$	<i>mumu-ga zaza-ga yoyo-ga kikotta-to hihatta-to tet<u>atta</u></i>
	$[N_3[N_2[N_1 V_1]V_2^*]V_3]$	<i>dodo-ga rara-ga mumu-ga hihuru-to tet<u>uru</u>-to kikoru</i>
	$[N_3[N_2[N_1 V_1]V_2^*]V_3^*]$	<i>dodo-ga mumu-ga zaza-ga tetaru-to ses<u>oru</u>-to hih<u>uru</u></i>
	$[N_3[N_2[N_1 V_1^*]V_2]V_3^*]$	<i>rara-ga mumu-ga yoyo-ga hih<u>aru</u>-to teturu-to ses<u>oru</u></i>
Simple _(L)	$[((NN)N_1) ((VV_1)V_1^*)]$	<i>rara-no gugu-no yoyo-ga kiki sesotte tet<u>uru</u></i>
	$[((NN)N_1) ((VV_1^*)V_1)]$	<i>gugu-no zaza-no dodo-ga kiki tet<u>atte</u> sesoru</i>
	$[((NN)N_1) ((VV_1^*)V_1^*)]$	<i>yoyo-no rara-no mumu-ga tete hih<u>otte</u> kik<u>otta</u></i>
Conjoined _(L)	$[N_1 V_1][N_2 V_2][N_3 V_3^*]$	<i>dodo-ga tetotte mumu-ga sesutte zaza-ga hih<u>oru</u></i>
	$[N_1 V_1][N_2 V_2^*][N_3 V_3]$	<i>gugu-ga kikutte zaza-ga tet<u>otte</u> dodo-ga sesoru</i>
	$[N_1 V_1][N_2 V_2^*][N_3 V_3^*]$	<i>zaza-ga sesatte yoyo-ga kik<u>utte</u> gugu-ga tet<u>otta</u></i>
	$[N_1 V_1^*][N_2 V_2][N_3 V_3^*]$	<i>mumu-ga ses<u>otte</u> rara-ga kikatte dodo-ga hih<u>utta</u></i>
Reverse _(L)	$A_3 A_2 A_1 B_1 B_2 B_3^*$	<i>gazaza rusose gunogu gunogu rusose <u>nozaza</u></i>
	$A_3 A_2 A_1 B_1 B_2^* B_3^*$	<i>ragara hiruhu gayoyo gayoyo <u>huruhi gazaza</u></i>
	$A_3 A_2 A_1 B_1 B_2^* B_3$	<i>serusa gugagu hohiru hohiru <u>noququ</u> serusa</i>
	$A_3 A_2 A_1 B_1^* B_2 B_3^*$	<i>gunogu ruteta nododo <u>noyoyo</u> ruteta <u>gugagu</u></i>
Same _(L)	$A_1 A_2 A_3 B_1 B_2 B_3^*$	<i>hiruho gunogu haruhi hiruho gunogu <u>hahiru</u></i>
	$A_1 A_2 A_3 B_1 B_2^* B_3^*$	<i>dogado rutetu zagaza dogado <u>suruse zanoza</u></i>
	$A_1 A_2 A_3 B_1 B_2^* B_3$	<i>kattaki yonoyo tutetta kattaki <u>nododo</u> tutetta</i>
	$A_1 A_2 A_3 B_1^* B_2 B_3^*$	<i>noyoyo tahihha munomu <u>nododo</u> tahihha <u>mugamu</u></i>

For the Simple_(L), there were also other error types, i.e., $V_1^*(VV_1^*)$, $V_1(VV_1^*)$, and $V_1^*(VV_1)$. For the Nested_(L) and Conjoined_(L), we included nonmatching stimuli with a maximum variety of vowels in Vs, as shown here; for the Simple_(L), we included nonmatching stimuli with a least variety of vowels in Vs. Therefore, the strategy of noting the variety of vowels was not effective. Under all conditions, there were more variations in error for the long than short stimuli.

Syntactic Computation in the Human Brain: The Degree of Merger as a Key Factor

Shinri Ohta, Naoki Fukui, Kuniyoshi L. Sakai

Appendix S1 Theoretical issues

Theoretical Background

Sentences with different constructions have been previously discussed in terms of the acceptability of sentences [cf. [1], p. 12].

- (i) nested constructions,
- (ii) self-embedded constructions,
- (iii) multiple-branching constructions,
- (iv) left-branching constructions, and
- (v) right-branching constructions

The nested constructions are created by *centrally* embedding a phrase within another phrase (with some non-null element to its left and some non-null element to its right), and the self-embedded constructions are the special case of nested constructions when nesting occurs within the *same* type of phrases (e.g., noun phrases). The multiple-branching constructions are made by conjoining phrases at the same hierarchical level, and the left/right-branching constructions are yielded by merging a phrase in the left-most or right-most phrase. The degrees of nesting and self-embedding have been already proposed to model the understanding of sentences [2]. By generalizing this attractive idea to include any constructions with merged phrases, we introduced DoM as a key computational concept.

We have proposed that various other “miscellaneous” operations that have been employed in the linguistics literature, such as Agree, Scope determination, Copy, etc., are in fact different manifestations of one and the same, i.e., more generalized, operation of Search [3]. Thus, Agree, which has been assumed to be an operation mainly responsible for the agreement of grammatical features, is actually not an operation specific to agreement, but rather is just an instance of the basic operation Search, when it assigns specific features. Human language, therefore, should minimally contain only two universal operations, Merge and Search.

As regards the formal symbol sequences beyond the bounds of finite state languages, three specific types of language have been discussed in the linguistic literature: (i) “counter language”, (ii) “mirror-image language”, and (iii) “copying language” [cf. [4], p. 21].

- (i) $ab, aabb, aaabbb, \dots$, and in general, all sentences consisting of n occurrences of a followed by n occurrences of b and only these;
- (ii) $aa, bb, abba, baab, aaaa, bbbb, aabbaa, abbbba, \dots$, and in general, all sentences consisting of a string X followed by the ‘mirror image’ of X (i.e., X in reverse), and only these;
- (iii) $aa, bb, abab, baba, aaaa, bbbb, aabaab, abbabb, \dots$, and in general, all sentences consisting of a string X of a ’s and b ’s followed by the identical string X , and only these.

The counter language can be handled by a counting mechanism to match the number of each symbol, whereas the mirror-image language contains a mirror-image dependency, requiring more than a mere counter. If the number of symbols is not fixed (i.e., infinite), both of these languages are to be generated by context-free (simple) phrase structure grammars, while the copying language

with a cross-serial dependency clearly goes beyond the bounds of context-free phrase structure grammars, requiring a more powerful device, viz., context-sensitive phrase structure grammars or transformational grammars [5,6]. In the present study, the Reverse and Same under the string conditions took the above type (i) of $A^n B^n$. As regards the matching orders, the Reverse took the type (ii) of $A_2 A_1 B_1 B_2$ or $A_3 A_2 A_1 B_1 B_2 B_3$, while the Same took the type (iii) of $A_1 A_2 B_1 B_2$ or $A_1 A_2 A_3 B_1 B_2 B_3$.

Because the number of symbols is inevitably fixed (i.e., finite) in any actual experiments, it should be noted that any symbol sequences can be expressed by a regular (finite state) grammar, i.e., the least powerful grammar in the so-called Chomsky hierarchy. Therefore, one cannot claim in principle that individual grammars (e.g., context-free phrase structure vs. regular grammars) are differentially represented in the brain, and the neural representation of individual grammars was *not* within the scope of the present study. Besides various models examined, other nonstructural and nonsymbolic models with simple recurrent networks have been proposed to process some examples of even context-free and context-sensitive phrase structure languages, generalizing to some degree to longer strings than the training set [7]. However, these models do not account for any parametric modulation of activations in the present study, except the length of sentences.

Operational Definitions of All Factors Examined

We operationally defined syntactic factors within an entire sentence (see Table 1) as follows. If a tree structure [a Phrase-marker (P-marker) associated with a linguistic expression] contains as its subtree a domain in which a node N immediately dominates n elements ($n > 1$), then we can say that the domain constitutes a *merged* structure. Note that under the binary Merge hypothesis, n equals 2, except for relatively rare “multiple branching” structures (see Appendix S2). In the present study, we abstract away from the noun/verb (N/V or \bar{N}/\bar{V}) vs. noun phrase/verb phrase (NP/VP) distinction, as well as the sentence (S) vs. complementizer phrase (CP or \bar{S}) distinction. The operational definitions of syntactic factors examined here are as follows (see Figures 1A and 2A). “Number of Merge” is the total number of binary branches. “Number of Search” is the total number of correspondences between sample and comparison stimuli. DoM is the largest integer m meeting the following condition: there is a continuous path passing through $m + 1$ nodes N_0, \dots, N_m , where each N_i ($i \geq 1$) is *merged* in the subtree dominated by N_{i-1} .

From both theoretical and experimental points of view, we also examined in detail the validity of other structure-based models, here categorized as “other linguistic factors”. If a merged structure is surrounded by non-null elements on both sides, we get a “nested” structure. If a nesting structure occurs within the *same* type of elements, the structure is called a “self-embedded” structure. “Degree of nesting” (or “degree of self-embedding”) is the largest integer m meeting the following condition (Figure S1A): there is a continuous path passing through $m + 1$ nodes N_0, \dots, N_m , where each N_i ($i \geq 1$) is nested (or fully self-embedded) in the subtree dominated by N_{i-1} [2]. “Number of nodes” is the total number of nonterminal nodes and terminal nodes (Figure S1B).

As regards “depth of postponed symbols”, its original definition was for *producing* a given output sequence [8]. As we tested stimuli for *understanding* a given input sequence, we reversed the numbering for the listener/reader as follows. “Depth of postponed symbols” is the amount of temporary storage needed for parsing a given input sequence, which can be calculated in the following way: first, number the branches of each node from 0 to $n - 1$ (Figure S2, the digits shown in red), where n is the number of branches from that node. Start numbering from the left. Then, compute the depth d of each terminal node by adding together the numbers written along all branches leading to that terminal node, starting from the leftmost branch (i.e., the first input for the listener/reader).

According to the dependency locality theory [9], two components of sentence parsing that consume computational resources have been proposed: “integration costs”, which are connecting words into the structure for the input thus far, and “storage costs”, which are the minimum number of words required to complete the current input as a grammatical sentence (Figure S3A). By adding

together “new discourse referents” (Ns and Vs in our paradigm) and “structural integrations” (the number of discourse referents in the intervening region), “integration costs” are obtained. According to the similarity-based interference theory [10], there are combined effects of “retroactive interference” (the number of nominative NPs between the subject-verb pair, when a verb is processed) and “proactive interference” (the number of nominative NPs, which are prior to the subject and still active in the parse) on syntactic attachments (Figure S3B). By adding both interference effects together, “syntactic interference” is obtained. Another source of interference is “positional similarity”, which is the number of adjacent syntactically similar NPs (i.e., marked with similar case markers).

Nonlinguistic factors may also variably contribute to the processing load of sentences with different constructions. At least three basic nonlinguistic factors may be involved in our experiment: memory span, counting, and “number of encoding”. The operational definitions of these nonlinguistic factors are as follows. Memory span is the maximum cost needed to maintain an item for matching against intervening or skipped stimuli (e.g., zero for $N_1 V_1$ in the Conjoined_(S), and one for $N_1 VV_1$ in the Simple_(S)), and its operational definition is the maximum number of cusps in the curved arrows (Figures 1 and 2). Counting is an operation needed to track symbol repetition, and its operational definition is the maximum number of consecutively repeated symbols (e.g., zero for NVNV or NVNVNV, one for NNVV, and two for NNNVVV). Encoding is memorization of features necessary for matching, and “number of encoding” is the total number of sample and comparison stimuli. Memory span and counting were considered temporal order-related factors in our experiment; memory span was related to matching orders, while counting was related to symbol orders. On the other hand, memory span and “number of encoding” were memory-related factors.

References

1. Chomsky N (1965) *Aspects of the Theory of Syntax*. Cambridge, MA: The MIT Press. 251 p.
2. Miller GA, Chomsky N (1963) Finitary Models of Language Users. In: Luce RD, Bush RR, Galanter E, editors. *Handbook of Mathematical Psychology*. New York, NY: John Wiley and Sons. pp. 419-491.
3. Fukui N, Sakai H (2003) The visibility guideline for functional categories: Verb raising in Japanese and related issues. *Lingua* 113: 321-375.
4. Chomsky N (1957) *Syntactic Structures*. The Hague: Mouton Publishers. 117 p.
5. Chomsky N (1959) On certain formal properties of grammars. *Inform Control* 2: 137-167.
6. Hopcroft JE, Ullman JD (1979) *Introduction to Automata Theory, Languages, and Computation*. Reading, MA: Addison-Wesley. 418 p.
7. Rodriguez P (2001) Simple recurrent networks learn context-free and context-sensitive languages by counting. *Neural Comput* 13: 2093-2118.
8. Yngve VH (1960) A model and an hypothesis for language structure. *Proc Am Philos Soc* 104: 444-466.
9. Gibson E (2000) The Dependency Locality Theory: A Distance-Based Theory of Linguistic Complexity. In: Marantz A, Miyashita Y, O'Neil W, editors. *Image, Language, Brain: Papers from the First Mind Articulation Project Symposium*. Cambridge, MA: The MIT Press. pp. 95-126.
10. Lewis R, Nakayama M (2002) Syntactic and positional similarity effects in the processing of Japanese embeddings. In: Nakayama M, editor. *Sentence Processing in East Asian Languages*. Stanford, CA: CSLI Publications. pp. 85-111.

Syntactic Computation in the Human Brain:

The Degree of Merger as a Key Factor

Shinri Ohta, Naoki Fukui, Kuniyoshi L. Sakai

Appendix S2 Detailed information about the stimuli

In the Japanese language, all regular verbs are either of *ichidan* (one-tier) or *godan* (five-tier) conjugation [1]. In our experiments, Vs took a nonpast-tense form (-*ru*), past-tense form (-*ta*), or gerundive form (-*te*) (e.g., “*teteru*”, “*teteta*”, and “*tetete*”), inflecting like normal *ichidan* verbs (e.g., “*tateru*”, “*tateta*”, and “*tatete*”: “*build*”, “*built*”, and “*building*”). In a subordinate clause, Vs took tense markers (-*ru*/*-ta*) just before -*to*. The last V in an entire sentence always ended with tense markers. As the tense of verbs can be independently marked in a Japanese nested sentence (e.g., “*Taro-ga Hanako-ga utau-to omotta*”: “*Taro thought that Hanako would sing*”), we regarded “number of tense markers (-*ru*/*-ta*)” as one of other linguistic factors (Table 1).

In Japanese, there are a number of morphologically/semantically related pairs of transitive and intransitive verbs with vowel changes (e.g., “*kakeru*” and “*kakaru*”: “*hang up*” and “*hang down*”) [2]. Not only such vowel dissimilation, but vowel assimilation, i.e., *vowel harmony*, is commonly observed in natural languages [3]. Vowel harmony is possible between adjacent or distant vowels within a word [4]. Indeed, the presence of vowel harmony has been indicated in the history of the Japanese language [1]. In the present study, vowel harmony was adopted to change the last, i.e., second, vowel of the verb root, so that this vowel harmonized with the vowel (i.e., /a/, /u/, or /o/) of the corresponding subject (e.g., “*rara-ga tetaru*” and “*rara-ga tetatta*” from “*teteru*” and “*teteta*”, respectively). This property of vowel harmony made each V_i inflect like a normal *godan* verb, which root always ends in a consonant, and this change made V_i distinct from the original form of *ichidan* verbs. It should be noted that V_i of nonmatching stimuli also inflected like normal *godan* verbs (Tables S1 and S2). When a *godan* verb root ends in *r*, *t*, or *w*, an euphonic change of a geminated consonant *t* occurs before -*ta* or -*te* (e.g., “*kakatta*” from “*kakari-ta*”: “*hung down*”). Each V_i with -*ta* or -*te* (e.g., “*tetatta*”) was consistent with its adverbial form *ri* (“*tetari-ta*”), *ti* (“*tetati-ta*”), or (*w*)*i* (“*tetai-ta*”). On the other hand, the first verb of a compound verb remained an *ichidan* verb in an adverbial form (e.g., “*tete*”).

There are some grammatical, but noncanonical (i.e., in a special use), usages of -*ga* or -*no*: parallel subjects marked with -*ga* (e.g., “*Taro-ga yuujin-ga sorezore utatta*”: “*Taro and his friend each sang*”), an object marked with -*ga* (e.g., “*Taro-ga yuujin-ga suki-da*”: “*Taro likes his friend*”), an external possessor marked with -*ga* (e.g., “*Taro-ga yuujin-ga sinsetu-da*”: “*Taro’s friend is kind*”), and a subject marked with -*no* (e.g., “*Taro-no suki-na yuujin*”: “*the friend Taro likes*”). Considering such canonicity, we regarded “number of case markers (-*ga*/*-no*)” as one of other linguistic factors (Table 1). However, these noncanonical case markings are rare in both comprehension and production, as shown by previous behavioral experiments [5,6]. We assured the participants that case markings for our stimuli were always canonical (see Appendix S3). Actual usage of canonical case markings was fully guaranteed by the high accuracy under the sentence conditions (Figure 3A), as the matching task could not be performed correctly if such noncanonical case markings were employed.

We imposed three following constraints for the letter strings. First, for the first half of a string (\mathcal{A}_s), letter strings derived from Ns and Vs (denoted here as \mathcal{N}_s and \mathcal{V}_s , respectively) were in the order of $\mathcal{N}\mathcal{V}$ or $\mathcal{V}\mathcal{N}$ for the short stimuli, and $\mathcal{N}\mathcal{V}\mathcal{N}$ or $\mathcal{V}\mathcal{N}\mathcal{V}$ for the long stimuli. Secondly, we avoided endings with -*ga*/*-no* for \mathcal{N}_s , but some \mathcal{V}_s with -*ru*/*-ta* endings were used as stimuli. Lastly, neither Vs with -*to*/*-te* endings nor Vs in the adverbial form were used for making \mathcal{V}_s .

Examples of long sentences are shown in Figure 2A. For the Nested_(L), a sentence at the lowest hierarchical level (S) was self-embedded twice into an entire sentence (S’). For the Simple_(L), we tested both stimuli of (VV₁)V₁ and V₁(VV₁), where two verbs, i.e., VV₁ (a compound verb) and V₁, were conjoined. As regards the Ns in the Simple_(L), branching constructions were ambiguous between (NN)N₁ and N(NN₁), like “*Japanese history teacher*” ([[*Japanese history*] *teacher*] and [*Japanese* [*history teacher*]], respectively), both of which yielded the same DoM. For Japanese relative clauses, as well as for noun phrases with a genitive case marker *-no*, left-branching constructions predominate [cf. [7], pp.471-472]. For the Conjoined_(L), DoM for a node was increased by one from the highest nodes of an entire sentence (S’) to the same hierarchical level of conjoined sentences (Ss). Since all the conjoined sentences were equivalent with respect to their status in the multiple branching, two Merge applications involved here were assumed to follow, as a marked (i.e., “exceptional”) option, the associative law (i.e., [[*a # b*] # *c*] = [*a* # [*b* # *c*]], where # represents Merge). This type of associative Merge (which yields *n*-ary structures by applying *n* – 1 times) has been argued to be permissible for a certain class of marked constructions in human language [8]. In this way, the number of Merge in a sentence becomes always one less than the number of terminal nodes, *irrespective of sentence structures*.

References

1. Shibatani M (1990) *The Languages of Japan*. Cambridge, UK: Cambridge University Press. 411 p.
2. Tsujimura N (2007) *An Introduction to Japanese Linguistics*, 2nd Edition. Malden, MA: Blackwell Publishing. 501 p.
3. Nevins A (2010) *Locality in Vowel Harmony*. Cambridge, MA: The MIT Press. 244 p.
4. Mailhot F, Reiss C (2007) Computing long-distance dependencies in vowel harmony. *Biolinguistics* 1: 28-48.
5. Uehara K, Bradley DC (2002) Center-embedding Problem and the Contribution of Nominative Case Repetition. In: Nakayama M, editor. *Sentence Processing in East Asian Languages*. Stanford, CA: CSLI Publications. pp. 257-287.
6. Miyamoto ET (2002) Case markers as clause boundary inducers in Japanese. *J Psycholinguist Res* 31: 307-347.
7. Miller GA, Chomsky N (1963) Finitary Models of Language Users. In: Luce RD, Bush RR, Galanter E, editors. *Handbook of Mathematical Psychology*. New York, NY: John Wiley and Sons. pp. 419-491.
8. Fukui N (2011) Merge and Bare Phrase Structure. In: Boeckx C, editor. *The Oxford Handbook of Linguistic Minimalism*. Oxford, UK: Oxford University Press. pp. 73-95.

Syntactic Computation in the Human Brain: The Degree of Merger as a Key Factor

Shinri Ohta, Naoki Fukui, Kuniyoshi L. Sakai

Appendix S3 Task instructions and training procedures

Task Instructions

Before the experiments, all participants were fully informed about the stimuli and task. We instructed and trained the sentence conditions in the order of Simple_(S), Conjoined_(S), Nested_(S), Simple_(L), Conjoined_(L), and Nested_(L), as the number of vowel extraction increased in this order. The following is a translation of task instructions in Japanese.

[Day 1 instructions]

Words used in a task

Special pseudowords and letter strings will be used in a task. **You don't have to remember the following words.**

Pseudowords used as nouns (six kinds) are: “rara”, “zaza”, “mumu”, “gugu”, “yoyo”, and “dodo”.

Grammatical particles attached to the noun (two kinds) are: “-ga”, which marks a subject, and “-no”, which marks a modifier (e.g., “zaza-ga”, “mumu-no”, etc.).

Pseudowords used as verbs (four kinds) are: “hihi”, “kiki”, “sese”, and “tete”.

Verb endings (five kinds) and verb conjugation patterns are:

1. Verb endings: “-ru” (present tense), “-tta” (past tense), “-tte” (e.g., “utatte”), “-ru-to” (e.g., “odoru-to”), and “-tta-to” (e.g., “odotta-to”); and
2. Verb conjugation: **When a subject corresponds to a verb within a sentence, the latter vowel (second syllable) of the verb root will change in order to match the vowel of the corresponding subject.** You don't have to recall the root form of the verb.

Correct examples:

“rara-ga hiharu” (conjugation of “hihi”)
“mumu-ga kikutta” (conjugation of “kiki”)
“dodo-ga sesoru” (conjugation of “sese”)
“rara-ga tetatta” (conjugation of “tete”)

Wrong examples:

“zaza-ga hihoru”
“gugu-ga kikotta”
“yoyo-ga sesuru”
“zaza-ga tetoru”

Other letter strings are: “ragara”, “nogugu”, “huhhita”, “kottaki”, etc.

Types of conditions

- 1) Grammar (G) conditions, 2) Memory (M) condition, and 3) Reverse (R) condition.

You will perform a task under three types of conditions during scanning of your head images inside the MRI scanner. At the beginning of each trial, a cue (“G”, “M”, or “R”) denoting a condition will appear. By noting these cues, please judge which condition will be presented. During the task, a small red cross remains at the center of the monitor. Fixate this red cross as much as you can, but you may blink. *Please do not speak or read aloud during the task.*

During the task, please respond by pressing a button. Hold a switch box by both hands, with a red button to your right. When you press one, use your right thumb without looking at the buttons. Please note the following crucial points:

- Press a button only once, when necessary.
- Please press the button as fast as you can, while the stimuli are presented.
- Every stimulus is consisted of a cue denoting a condition and of four or six phrases or strings.
- Do not press a button before the last stimulus.

1. Grammar conditions

There are six patterns under the Grammar conditions. The task is to judge whether or not the vowel of a subject is matched with the last vowel of the corresponding verb root. At the beginning of each trial, the cue “G” denoting a condition will appear. Do not overlook this cue, so that you can answer correctly. For every pattern shown below, “G” will appear.

1.1 Pattern 1 (Simple_(S))

Pattern 1 is similar to “Taro-no ani-ga sinobi aruita” (“Taro’s brother sneaked around”). The subject is “ani”, and the verb is “aruita”. Please judge whether or not the vowel of the *second* phrase (subject) is matched with that of the *fourth* phrase (corresponding verb). Please press the *rightmost button* if correct, and press the *second one from the right* if wrong. Press the button quickly with your right thumb, while the fourth stimulus is presented.

Correct examples:

“zaza-no yoyo-ga hihi tetoru”

“zaza-no yoyo-ga hihi tetotta”

Wrong examples:

“zaza-no yoyo-ga tete sesuru”

“zaza-no yoyo-ga tete sesatta”

1.2 Pattern 2 (Conjoined_(S))

Pattern 2 is similar to “Taro-ga utatte Hanako-ga odoru” (“Taro sings, and Hanako dances”). Please judge whether or not the vowel of the *first* phrase is matched with that of the *second* phrase, and the vowel of the *third* phrase with that of the *fourth* phrase.

Correct examples:

“zaza-ga tetatte yoyo-ga kikoruru”

“zaza-ga tetatte yoyo-ga kikotta”

Wrong examples:

“zaza-ga tetatte yoyo-ga kikuru”

“zaza-ga tetotte yoyo-ga kikatta”

1.3 Pattern 3 (Nested_(S))

Pattern 3 is similar to “Taro-ga Hanako-ga utau-to omotta” (“Taro thought that Hanako would sing”). Please judge whether or not the vowel of the *first* phrase is matched with that of the *fourth* phrase, and the vowel of the *second* phrase with that of the *third* phrase.

Correct examples:

“zaza-ga yoyo-ga tetoru-to kikaru”

“zaza-ga yoyo-ga tetotta-to kikatta”

Wrong examples:

“zaza-ga yoyo-ga tetoru-to kikoruru”

“zaza-ga yoyo-ga tetatta-to kikutta”

Tips

1. Please do not forget the first noun.
2. There are sentences that have multiple errors; *please check them up to the last phrase*.
3. Please press the button *as fast as you can, while the last phrase is presented*.
4. *When the last phrase disappears, please quit the button press, and concentrate on the next trial.*

2. Memory condition (Same_(S))

The task is to memorize the presented letter strings. At the beginning of each trial, the cue “M” denoting a condition will appear. Do not overlook this cue, so that you can answer correctly.

Four strings will appear one by one on the monitor. Please memorize these four strings, and judge whether or not the *first* and *third* strings, and the *second* and *fourth* ones, are exactly the same.

Correct examples:

“yogayo tarute yogayo tarute”
“hiruhu garara hiruhu garara”

Wrong examples:

“yonoyo tettata yonoyo tatetta”
“hihata nogugu settaso gunogu”

3. Reverse condition (Reverse_(S))

The task is to memorize the presented letter strings. At the beginning of each trial, the cue “R” denoting a condition will appear. Do not overlook this cue, so that you can answer correctly.

Four strings will appear one by one on the monitor. Please memorize these four strings, and judge whether or not the *first* and *fourth* strings, and the *second* and *third* ones, are exactly the same.

Correct examples:

“yogayo tarute tarute yogayo”
“hiruhu garara garara hiruhu”

Wrong examples:

“yonoyo tettata tatetta yonoyo”
“hihata nogugu nogugu settaso”

[Day 2 instructions]

1. Grammar conditions

1.1 Pattern 4 (Simple_(L))

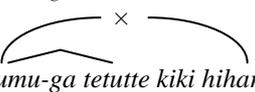
Pattern 4 is similar to “Taro-no ani-no yujin-ga hasiri mawatte utatta” (“Taro’s brother’s friend run around and sang”). The subject is “yujin”, and the verbs are “mawatte” and “utatta”. Please judge whether or not the vowel of the *third* phrase is matched with that of the *fourth* or *fifth* phrase, and the vowel of the *third* phrase with that of the *sixth* phrase. Please press *the rightmost button* if correct, and press *the second one from the right* if wrong. Press the button quickly with your right thumb, while the sixth stimulus is presented.

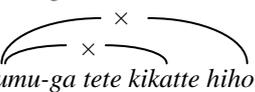
Correct examples:

“zaza-no yoyo-no mumu-ga tete kikutte hihuru”


“zaza-no yoyo-no mumu-ga tetutte kiki hihutta”


Wrong examples:

“zaza-no yoyo-no mumu-ga tetutte kiki hiharu”


“zaza-no yoyo-no mumu-ga tete kikatte hihotta”


1.2 Pattern 5 (Conjoined_(L))

Pattern 5 is similar to “Taro-ga odotte Hanako-ga utatte Jiro-ga asobu” (“Taro dances, Hanako sings, and Jiro plays”). Please judge whether or not the vowel of the *first* phrase is matched with that of the *second* phrase, the vowel of the *third* phrase with that of the *fourth* phrase, and the vowel of the *fifth* phrase with that of the *sixth* phrase.

Correct examples:

“zaza-ga tetatte yoyo-ga kikotte mumu-ga hihuru”

Wrong examples:

“zaza-ga tetatte yoyo-ga kikotte mumu-ga hihutta”

“zaza-ga tetatte yoyo-ga kikotte mumu-ga hiharu”

“zaza-ga tetutte yoyo-ga kikotte mumu-ga hihatta”

1.3 Pattern 6 (Nested_(L))

Pattern 6 is similar to “Taro-ga Hanako-ga Jiro-ga utau-to omou-to kangaeta” (“Taro supposed that Hanako would think that Jiro would sing”). Please judge whether or not the vowel of the *first* phrase is matched with that of the *sixth* phrase, the vowel of the *second* phrase with that of the *fifth* phrase, and the vowel of the *third* phrase with that of the *fourth* phrase.

Correct examples:

“zaza-ga yoyo-ga mumu-ga teturu-to kikoru-to hiharu”

“zaza-ga yoyo-ga mumu-ga tetutta-to kikotta-to hihatta”

Wrong examples:

“zaza-ga yoyo-ga mumu-ga teturu-to kikoru-to hihoru”

“zaza-ga yoyo-ga mumu-ga tetutta-to kikatta-to hihotta”

2. Memory condition (Same_(L))

Six strings will appear one by one on the monitor. Please memorize these six strings, and judge whether or not the *first* and *fourth* strings, the *second* and *fifth* ones, and the *third* and *sixth* ones, are exactly the same.

Correct examples:

“yogayo teruta gagugu yogayo teruta gagugu”

“huruhi garara ruseso huruhi garara ruseso”

Wrong examples:

“noyoyo tattate gunogu noyoyo tattate munomu”

“huttahi ranora sottase hihhuta ranora tattate”

3. Reverse condition (Reverse_(L))

Six strings will appear one by one on the monitor. Please memorize these six strings, and judge whether or not the *first* and *sixth* strings, the *second* and *fifth* ones, and the *third* and *fourth* ones, are exactly the same.

Correct examples:

“yogayo teruta gagugu gagugu teruta yogayo”

“huruhi garara ruseso ruseso garara huruhi”

Wrong examples:

“noyoyo tattate gunogu munomu tattate noyoyo”

“huttahi ranora sottase tattate ranora hihhuta”

Training Procedures

Experiments with short or long stimuli were performed on separate days in the order short, then long. Before scanning, the participants were trained until they scored 80 % at each of the following stages. Each of the Simple, Conjoined, and Nested were separately tested with self-paced reading in 10 to 30 trials with this order. Only one participant took 50 trials for the Nested_(L). When

these conditions were randomized, all participants took additional 20 or 40 trials. The Reverse and Same were much easier to perform, and so required only 10 or 20 trials for a randomized sequence from the beginning. When all of these conditions were mixed, participants additionally took 20 or 40 trials. Finally, participants were tested on a sequence of all conditions at the regular rate of presentation (see Figure 1C and 1D) in 40 and 20 trials for the short and long stimuli, respectively. Only three participants required 60 or 80 trials for the short stimuli. Due to technical problems, four participants were retested with short or long stimuli on another day; two participants received 20 trials for the mixed conditions with self-paced reading, and all received 20–60 trials for the final stage.

Syntactic Computation in the Human Brain: The Degree of Merger as a Key Factor

Shinri Ohta, Naoki Fukui, Kuniyoshi L. Sakai

Appendix S4 MRI data analyses

fMRI Data Analyses

Data analyses with fMRI were performed using SPM5 statistical parametric mapping software (Wellcome Department of Cognitive Neurology, London, UK; <http://www.fil.ion.ucl.ac.uk/spm/>) [1] on MATLAB software (Math Works, Natick, MA). The acquisition timing of each slice was corrected using the middle slice as a reference for the EPI data. We realigned the EPI data in multiple runs to the first volume in multiple runs, and removed runs that included data with a translation of > 2 mm in one of the three directions and with a rotation of $> 1.4^\circ$ around one of the three axes. Each participant's T1-weighted image was coregistered to the mean image of its own EPI data, followed by the segmentation of the T1-weighted image into grey- and white-matter probability maps. With this latter step, each individual brain was spatially normalized to the standard brain space as defined by Montreal Neurological Institute (MNI). Resampling was performed every 3 mm using seventh-degree B-spline interpolation, and images were smoothed with an isotropic Gaussian kernel of 9 mm full width at half maximum (FWHM).

Using these preprocessed data, low-frequency noise and global changes in activity were further removed (cut-off period: 128 s). Hemodynamic responses induced by the trials were modeled with a box-car function with a duration of 3.5 s (short stimuli) or 5.1 s (long stimuli), i.e., from the onset of the first stimulus (N or A) to 500 ms after the disappearance of the last stimulus (Figure 1C and 1D), and the box-car function was convolved with a hemodynamic function that peaked at 6 s. These functions were used as run-specific covariates for matching or nonmatching stimuli for each condition in a general linear model (GLM). Only event-related responses to correct trials were analyzed.

For random effects analyses using either an analysis of covariance (ANCOVA) with t -statistics or a one-sample t -test, contrast images were generated for each participant and used for intersubject comparisons. For all fMRI data analyses, the statistical threshold was set to $P < 0.05$ for the voxel level, corrected for multiple comparisons [family-wise error (FWE) correction] across the whole brain. To exclude any general cognitive factors related to task difficulty from the evaluation of cortical activation, the accuracy was used as a nuisance variable for each contrast of random effects analyses. If a subtraction between conditions served as a reference, an exclusive mask (uncorrected $P < 0.01$) was applied to reduce the contribution of deactivation. For example, in Nested' $>$ Simple', an exclusive mask of the [Conjoined – Simple] contrast was applied to reduce the contribution of deactivation in Simple'.

We used a factorial design for a two-way ANCOVA with condition [Nested', Simple'] \times length [Long, Short], where activations in Nested' or Simple' for long or short sentences were estimated for each participant, and then the main contrast (as denoted by a greater-than symbol) was calculated for intersubject comparisons. A one-way ANCOVA was used for a direct comparison of Nested'_(L) $>$ Simple'_(S), and for Nested'' $>$ Reverse'' contrasting the sentence and string conditions. To ensure an independent statistical test in the latter contrast, the inclusive mask of Nested' $>$ Simple' (corrected $P < 0.05$) was iteratively applied to the contrast image of each participant, thereby leaving out that participant from the calculation of the mask, i.e., the “leave-one-subject-out cross-validation” approach [2]. One-sample t -test was used for [Nonmatching – Matching] separately under the sentence and string conditions, and for Reverse''.

For the anatomical identification of activated regions, we basically used the Anatomical Automatic Labeling method [3]. The percent signal changes averaged across the voxels in each activated region were extracted using the MarsBaR-toolbox (<http://marsbar.sourceforge.net/>). To fit a model of each factor to activations, a fitting scale and residual sum of squares (RSS) were calculated with MATLAB, and the fitted values were obtained by multiplying the estimates by the fitting scale. For a no-intercept model, coefficient of determination (r^2) should be calculated as $r^2 = 1 - \Sigma(y - \hat{y})^2 / \Sigma y^2$, where \hat{y} and y denote the fitted values and the signal changes for each contrast, respectively [4]. For this calculation, we used R software (<http://www.r-project.org/>). By using a restricted maximum-likelihood method, we further fitted “linear mixed-effects models” with individual activations as dependent variables, the estimates of each factor as a regressor, and the participants as random effects. For this calculation, we used an nlme (linear and nonlinear mixed-effects models) package (<http://cran.r-project.org/web/packages/nlme/>) on R software.

DCM Data Analyses

Data analyses with DCM were performed using DCM10 on SPM 8 [5]. For DCM analyses, we concatenated the scans from the separate runs, and reanalyzed the preprocessed data with GLM, which contained regressors representing the Nested, Simple, Reverse, and Same (correct trials alone; see Table 2), as well as a regressor representing all conditions (correct trials alone including Conjoined). In addition, the effects of transition between runs were taken into account by adding regressors for each run.

We imposed following functional and anatomical constraints for selecting ROIs (regions of interest) in L. F3op/F3t or L. SMG for each participant with the individually preprocessed data [6]. The individual local maxima should be significant in Nested” – Reverse” estimated for each participant (uncorrected $P < 0.05$), nearest to the group local maxima of Nested” > Reverse” within twice the FWHM of the smoothing kernel, and obviously in L. F3op/F3t or L. SMG. The averaged MNI coordinates of these individual local maxima were (–50, 26, 25) and (–39, –45, 43) for L. F3op/F3t and L. SMG, respectively. With the volume-of-interest (VOI) tool on SPM8, the time series was extracted by taking the first eigenvariate across all suprathreshold voxels within 6 mm of the individual local maxima.

We specified nine models with systematic variations in a modulatory effect and driving inputs (Figure S4). The regressor representing the Nested was used for a modulatory effect, whereas that representing all conditions was used for driving inputs. After estimating all models for each participant, we identified the most likely model by using random-effects Bayesian model selection (BMS) on DCM10. Inferences from BMS can be based on the expected probability, i.e., the expected likelihood of obtaining the model for any randomly selected participants, or on the exceedance probability, i.e., the probability that the model is a better fit to the data than any other models tested. We adopted the family inference method [7], in which three modulatory families were compared first, and then the input models within the winning family were further compared. After determining the best model, the parameter estimates of this particular model were evaluated by a one-sample t -test [8].

DTI Data Analyses

Data analyses with DTI were performed using FSL [Oxford Centre for Functional MRI of the Brain’s (FMRIB) Software Library 4.1.7; <http://fsl.fmrib.ox.ac.uk/fsl/fslwiki/>] and FDT (FMRIB’s Diffusion Toolbox 2.0) [9]. Diffusion-weighted images were first resliced to isotropic voxel of 1 mm^3 , and then eddy current distortions and motion artifacts were corrected using affine registration to the b_0 image. We then extracted the brain shape from the b_0 image, and created the binary mask image (i.e., zero for the outside of the brain) for each participant. Markov Chain Monte Carlo sampling was performed to build up distributions on diffusion parameters at each voxel, which allowed for an estimation of the most probable pathway by Bayesian estimation (number of fibers

modeled per voxel = 2) [10]. The implicit modeling of noise in a probabilistic model made a fiber tracking near grey matter possible.

By using FLIRT (FMRIB's Linear Image Registration Tool) on FSL, the b0 image was first coregistered to the individual T1-weighted image for each participant, and the T1-weighted image was spatially normalized to the MNI space by using both affine and nonlinear transformations with FLIRT and FNIRT (FMRIB's Nonlinear Image Registration Tool). With the transformation matrices and estimated deformation fields, the peak MNI coordinates of each local maximum were transformed to the individual b0 images, and a sphere of 6-mm radius centered at the new coordinates was defined as a seed mask for the probabilistic tractography. All fiber tracking was conducted in an individual DTI space. In order to find the connections between these regions, we employed a two-ROI approach with two seed masks, which repeatedly sampled tracts from one seed mask, and retained only those tracts that passed through the other seed mask. Probabilistic fiber tracking was initiated from all voxels within the seed masks to generate 10,000 streamline samples, with a step length of 0.5 mm, a maximum number of steps of 2,000, a curvature threshold of 0.2 ($\pm 78.5^\circ$), and a loopcheck option.

In the connectivity distributions obtained, each voxel value represented the total number of the streamline samples passing through that voxel. The connectivity probability maps were then created for each participant by dividing the connectivity distributions with a sum of waytotal values, i.e., the total number of generated tracts from one seed mask that have reached the other seed mask. This normalization approach allowed for a comparison of connectivity probability values across participants; note that the pattern of connectivity did not change by this scaling. To remove any spurious connections, pathways in individual participants were thresholded to include only voxels that had at least 1 % connectivity probability values [11]. Thresholded pathways in each participant were then normalized as above and binarized using "fslmaths" on FSL. The binarized pathways were overlaid across participants to produce population probability maps for each pathway, in which voxel values represent the number of participants with a pathway through that voxel. These population probability maps with thresholding (at least seven out of 15 participants) were smoothed and shown using MRICroN software (<http://www.mccauslandcenter.sc.edu/mricro/mricron/>).

References

1. Friston KJ, Holmes AP, Worsley KJ, Poline J-P, Frith CD et al. (1995) Statistical parametric maps in functional imaging: A general linear approach. *Hum Brain Mapp* 2: 189-210.
2. Esterman M, Tamber-Rosenau BJ, Chiu Y-C, Yantis S (2010) Avoiding non-independence in fMRI data analysis: Leave one subject out. *Neuroimage* 50: 572-576.
3. Tzourio-Mazoyer N, Landeau B, Papathanassiou D, Crivello F, Etard O et al. (2002) Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. *Neuroimage* 15: 273-289.
4. Kvålseth TO (1985) Cautionary note about R^2 . *Am Stat* 39: 279-285.
5. Friston KJ, Harrison L, Penny W (2003) Dynamic causal modelling. *Neuroimage* 19: 1273-1302.
6. Stephan KE, Marshall JC, Penny WD, Friston KJ, Fink GR (2007) Interhemispheric integration of visual processing during task-driven lateralization. *J Neurosci* 27: 3512-3522.
7. Penny WD, Stephan KE, Daunizeau J, Rosa MJ, Friston KJ et al. (2010) Comparing families of dynamic causal models. *PLoS Comput Biol* 6, e1000709: 1-14.
8. Stephan KE, Penny WD, Moran RJ, den Ouden HEM, Daunizeau J et al. (2010) Ten simple rules for dynamic causal modeling. *Neuroimage* 49: 3099-3109.
9. Smith SM, Jenkinson M, Woolrich MW, Beckmann CF, Behrens TEJ et al. (2004) Advances in functional and structural MR image analysis and implementation as FSL. *Neuroimage* 23: S208-S219.

10. Behrens TEJ, Berg HJ, Jbabdi S, Rushworth MFS, Woolrich MW (2007) Probabilistic diffusion tractography with multiple fibre orientations: What can we gain? *Neuroimage* 34: 144-155.
11. Flöel A, de Vries MH, Scholz J, Breitenstein C, Johansen-Berg H (2009) White matter integrity in the vicinity of Broca's area predicts grammar learning success. *Neuroimage* 47: 1974-1981.