

Cortical encoding of linguistic constituent with and without morphosyntactic cues

Claire H C Chang, Stanislas Dehaene, Denise H Wu, Wen-Jui Kuo, Christophe C Pallier

▶ **To cite this version:**

Claire H C Chang, Stanislas Dehaene, Denise H Wu, Wen-Jui Kuo, Christophe C Pallier. Cortical encoding of linguistic constituent with and without morphosyntactic cues. Cortex, 2020, 129, pp.281 - 295. 10.1016/j.cortex.2020.04.024 . hal-03025786

HAL Id: hal-03025786
https://hal.science/hal-03025786

Submitted on 26 Nov 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Cortical encoding of linguistic constituent with and without morphosyntactic cues

Claire H. C. Chang[1], Stanislas Dehaene[2, 3], Denise H. Wu[4], Wen-Jui Kuo[1, 5*], Christophe Pallier[2]

1. Institute of Neuroscience, National Yang-Ming University, Taipei, Taiwan
2. Cognitive Neuroimaging Unit, CEA DSV/I2BM, INSERM, Université Paris-Sud, Université Paris-Saclay, NeuroSpin Center, 91191 Gif-sur-Yvette, France
3. Collège de France, 11 Place Marcelin Berthelot, 75005 Paris, France
4. Institute of Cognitive Neuroscience, National Central University, Zhongli, Taiwan
5. Brain Research Center, National Yang-Ming University, Taipei, Taiwan

*To whom correspondence should be addressed: Wen-Jui Kuo, Institute of Neuroscience & Brain Research Center, National Yang-Ming University, Taipei, Taiwan (11221).

Email addresses: hcchang73@gmail.com (Claire H. C. Chang), *Stanislas*.*Dehaene*@cea.fr (Stanislas Dehaene), wuhsien@gmail.com (Denise H. Wu), wjkuo@ym.edu.tw (Wen-Jui Kuo), *christophe@pallier*.org (Christophe Pallier).

Declarations of interest: none

Abstract

This study examined the brain areas involved in combining words into larger units when there are few or no morphosyntactic cues. We manipulated constituent length in word strings of the same length under two conditions: Mandarin sentence, which had sparse morphosyntactic cues, and nominal phrase that had no morphosyntactic cues (e.g., [[honey mustard] [chicken burger]]). Contrasting sentences to word lists revealed a network that largely overlapped with the one reported in languages with rich morphosyntactic cues, including left IFGorb/IFGtri and areas along left STG/STS. Both conditions showed increased activation in left IFGtri/IFGorb in functional ROIs defined based on previous study in sentence processing, while the nominal phrases additionally revealed a constituent length effect in bilateral dorsal IFGtri, left IFGoper, left pMTG/pSTG, left IPL, and several subcortical areas, which might reflect an increased reliance on semantic and pragmatic information. Moreover, in upper left IFGtri/IFGoper and left thalamus/caudate, this effect increased with the participants' tendency to combine nouns into phrases. The absence of syntactic constraints on linguistic composition might highlight individual differences in cognitive control, which helps to integrate non-syntactic information.

Keywords: syntax; semantic; fMRI; Mandarin; context;

1. Introduction

One remarkable property of human languages is that they strongly rely on recursive combinatorial processes: phonemes combine into morphemes, which combine into words, which combine into phrases, which combine into sentences, thus forming a potentially infinite system. One worthy goal for neurolinguistics is to understand how such constituent structure is encoded at the neural level. While current brain imaging methods do not yet allow researchers to delve into the details of neural computations, they already permit to circumscribe the brain areas involved in this processing. In this spirit, using fMRI, Pallier, Devauchelle, & Dehaene (2011) sought the correlates of the representation of the linguistic constituent structure. They proposed a cumulative model of constituent building in which longer constituents are encoded by more complex neural assemblies, predicting stronger brain activation when constituent size increases. A paradigm was designed to verify this hypothesis, in which participants had to read strings of twelve words organized into syntactic constituents of increasing length (e.g., 2 constituents of 6 words or 1 constituent of 12 words). Conforming to the prediction of the proposed model, the fMRI signal increased with the size of constituents, within several regions of the left-lateralized language network, including the orbital part of left inferior frontal gyrus (IFGorb), the triangular par of IFG (IFGtri), temporal pole (TP), anterior superior temporal sulcus (aSTS), posterior STS (pSTS), and temporal-parietal junction (TPJ). The hierarchical constituent structure was further supported by magneto-encephalography and intracranial recording studies (Ding, Melloni, Zhang, Tian, & Poeppel, 2015; Nelson et al., 2017) showing the cortical tracking of constituent boundaries, namely, sudden neural response drops at the boundaries of multi-word constituents in language-related areas including STG/STS and IFG.

While Pallier et al. (2011) manipulated six levels of constituent length up to a full sentence, a meta-analysis of 19 recent imaging studies (9 in English, 4 in German, 3 in Dutch, 2 in Japanese, and 1 in French) that included the contrast between sentence vs. word list revealed an effect in the same set of regions (Zaccarella, Schell, & Friederici, 2017). Among them, several studies adopted jabberwocky stimuli, which were created by replacing content words in normal sentences with pseudowords while preserving

function words and suffixes, e.g., "*the couse that rits our treeve*." Jabberwocky sentence is a case in which constituent structure is built solely based on morphosyntactic cues. Studies comparing the structured jabberwocky sentences and the scrambled ones consistently showed an effect in left IFG and left pSTS (Goucha & Friederici, 2015; Matchin, Hammerly, & Lau, 2017; Pallier et al., 2011).

The majority of previous imaging studies that manipulated constituent length were performed in Indo-European languages. The current study examined constituent length effect in a language that is very different from the languages tested so far, i.e., Mandarin Chinese (Zhou, Ye, Cheung, & Chen, 2009). Historically, Mandarin is a member of the Sino-Tibetan family, which has been a distinct language family for at least 6000 years (Matisoff, 1991; Thurgood & LaPolla, 2003). Typologically, Indo-European languages, as well as Japanese, are categorized as synthetic languages, which have rich morphology. In contrast, one of the most striking features of Chinese is the relative sparseness of grammatical cues, including inflection (case, person, number, gender, etc.) and function words (of, the, etc.)(Greenberg, 1960; Li & Thompson, 1981). The finding that jabberwocky sentences still evoked a constituent length effect showed the importance of morphosyntactic cues in building constituent structure (Goucha & Friederici, 2015; Matchin et al., 2017; Pallier et al., 2011). Using Mandarin materials, this study examined whether the same constituent structure encoding system is recruited regardless of typological differences in the density of morphosyntactic cues.

Constituent structure can be built solely based on morphosyntactic cues, but it is not restricted to syntax and can be found even in non-linguistic domains, where there is no morphosyntactic cues in a strict sense, e.g. music (Dehaene, Meyniel, Wacongne, Wang, & Pallier, 2015; Hagoort, 2005; Jackendoff, 1999; Tai, 2005). Because morphosyntactic cues are sparse in Mandarin, it is hard to generate jabberwocky sentences with a dominant interpretation. Instead, we manipulated constituent lengths in nominal phrases that consisted of only nouns and contained no morphosyntactic cues (Table 1), e.g., 蜂蜜芥末雞肉漢堡, corresponding to "*honey mustard chicken burger*" and 暑期課程教材大綱, corresponding to *"summer course material layout."* These nominal phrases clearly have constituent structures (i.e. *[summer [course [material layout]]] and [[honey mustard][chicken burger]]*), which can be formed by generalized

rules. For example, "*summer*" can easily be replaced by words in a similar semantic category. Chinese sentences still contain morphosyntactic markers, although sparse, and have a basic order of subject-verb-objective. In contrast, there are no syntactic constraints on constituent structure in our nominal phrases. By comparing nominal phrase to sentence, we are able to examine whether the encoding of linguistic constituent structure under the nominal phrase condition engages similar brain regions regardless of the absence of morphosyntactic cues. Opposite to jabberwocky sentences, the constituent structure of these nominal phrases is built solely based on non-morphosyntactic information, e.g., semantic and pragmatic. That language parsing depends on the integration of different sources of information, whose weightings can vary between and within languages by their availability and reliability, is implemented and advocated in several language processing models (Bates & Macwhinney, 1989; Ellis, 2002; Gibson & Pearlmutter, 1998; Kuperberg, 2007; Martin, 2016; Seidenberg & MacDonald, 1999; Townsend, Bever, & Crocker, 2001). These nominal phrases are legal in Mandarin, though it is conceivable that individual variation in parsing might be large without the restriction from grammatical cues. Therefore, a constituent length assessment task was included after the fMRI scanning session

In brief, this study examined the encoding of constituent structure in a language with relatively sparse morphosyntactic cues, i.e., Mandarin Chinese. We included normal sentences, which still contain some morphosyntactic cues, and nominal phrases, which have no morphosyntactic cues. If a common substrate underlies constituent construction based on both morphosyntactic and non-morphosyntactic cues, it should show an effect of constituent length under both conditions. In regions where constituent construction is dependent on morphosyntactic cues, the effect should be found only under sentence condition. On the other hand, brain regions only involved in the nominal phrase condition are more likely to be engaged for semantic or pragmatic processing.

2. Materials and Methods

2.1. Participants

A total of twenty-three participants were tested. One was excluded because of excessive motion during scanning (range > 8 mm,). The nominal phrases condition was added after the first 5 participants were tested. Therefore, data for this condition were available in only 18 participants. The participants were students (age ranges from 18 to 29; 5 males) enrolled in various universities in Taipei, Taiwan. All were right-handed and native speakers of Mandarin. The protocol received the approval of the Institutional Review Board of National Yang-Ming University, and written consent was obtained before MR scanning

2.2. Stimuli

Two independent sets of stimuli were generated: one for the ‘sentence’ condition and one for the ‘nominal phrase’ condition. In the ‘sentence’ condition, the stimuli were sequences of words obtained by concatenating syntactic constituents or single words from sentences. In the ‘nominal phrase’ condition, the stimuli were sequences of four nouns forming nominal phrases of various lengths. For both types of stimuli, we manipulated constituent length at 3 levels (Table 1). Below we detail the methods of construction.

Table 1 Examples of Stimuli. Each word was displayed for 300 ms, with 1200 ms of the inter-stimulus interval.

Sentence condition

Constituent Length	Number of Constituent	Examples
Short 1	6	非常 提出 可以 父母 蛋糕 我們 (very) (propose) (can) (parent) (cake) (we)
Medium 3	2	提出 這個 方案 這些 原料 成本 (propose this plan) (those material cost) *propose this plan the cost of these materials*
Long 6	1	他們 可以 負擔 這些 原料 成本 (they can afford those material cost) *they can afford the cost of these materials.*

Nominal-phrase condition

Constituent Length	Number of Constituent	Examples
Short 1	4	漢堡 雞肉 芥末 蜂蜜 (burger) (chicken) (mustard) (honey)
Medium 2	2	雞肉 漢堡 蜂蜜 芥末 (chicken burger) (honey mustard)
Long 4	1	蜂蜜 芥末 雞肉 漢堡 (honey mustard chicken burger)

2.2.1. Sentence condition

120 sentences of 6 words were generated (Supplementary 1). All words were made up of two traditional Chinese characters. All sentences had a right-branching syntactic structure, starting with the subject. The first and last three words of the sentences were always proper syntactic constituents, corresponding to nominal phrase (NP), verb phrase (VP), or complementizer phrase (CP). Please see Spplplemtary 1 for the assignment of phrase type and Supplementary 2 for their distribution. We computed the

probability for the verbs at the third position to precede comma or period in the SINICA corpus 3.0 (http://asbc.iis.sinica.edu.tw/), so to verify that they are not obligatory transitive (Supplementary 3).

The NPs always ended with the head nouns, which were preceded by determiner, adjective, quantifier, or noun. All but one VPs began with the head verbs, which were followed by NPs as the object or CPs as the complement, sometimes with an adverb in between. One of the VPs began with an adverb. CPs always began with an NP as the subject and followed by verb + object, adverb + verb, modal verb + verb, or serial verbs.

These three-word-long constituents were extracted and recombined by pairs to form the stimuli of medium constituent length. Single words were extracted and recombined to form the stimuli of short constituent length. We manually filtered out the sequences where adjacent phrases (medium constituent) or words (short constituent) could combine and form longer constituents.

2.2.2. Nominal phrase condition

240 different nouns were used to generate 60 four-word phrases (Supplementary 1). They have left-branching structure with the head noun at the final position. Then, for each phrase, we shuffled the nouns in order to create a sequence of 2 noun-noun phrases and a sequence of 4 single nouns that did not form any phrase. Table 1 displays an example of a triplet of items obtained by this procedure.

The distinction between nominal phrase and compound word is often implicit in Mandarin because morphosyntactic markers, e.g. “of” in English, that help to identify phrases are sparse (Aikhenvald, 2007) and words are not separated by spaces in text. However, according to the criterions provided in the literature (Chao, 1968; Huang, 1984b; Packard, 2000), our stimuli conformed better to the definition of nominal phrase. First, our stimuli consisted of disyllable free words. None of them was bound morpheme. Second, the meaning of our stimuli is compositional instead of idiomatic. An example of English compound with idiomatic meaning is “ladybug.” Third, syntactic transformations are usually not applicable to word-internal structure, while it is applicable to our stimuli, e.g. conjunction reduction (蜂蜜芥末跟洋蔥芥末→蜂蜜跟洋蔥芥末; honey mustard and onion mustard→honey and onion mustard). In addition, we

computed the transition probabilities ($P(w_i \mid w_i\text{-}1)$) between all the neighboring words in the nominal phrase stimuli based on the SINICA corpus 3.0 (http://asbc.iis.sinica.edu.tw/). The mean transition probabilities are fairly low (short: 1%, medium 2%, long 2%). As a comparison, the transition probability between “suit” and “case,” which constitute the English compound “suitcase”, is 9% (based on Google Ngram in 2007, the latest year available. https://books.google.com/ngrams).

The construction of stimuli with different constituent lengths was done according to the intuitions of the first author and consulting colleagues who were native speakers of Mandarin. However, because we noticed that the intuitions varied from individuals to individuals, we asked each participant to judge the number of constituent in the string. More precisely, after they had completed the fMRI experiment, the participants were presented with all the stimuli previously displayed to them, and were asked verbally, “How many concatenated strings, or constituents, are there in this sequence? Please answer quickly without reflecting too much.” The instruction was presented visually before the assessment started: “請盯著螢幕上閃過的詞，回答你是否感覺這些字照順序串連起來? 全部串起來按 1，串成兩組按 2，串成三組按 3，完全不串按 4，請依直覺盡快做答。” The participants were free to provide any integer number. This was done both for the nominal phrase stimuli and the sentence stimuli. The length of the stimulus was divided by the reported number of constituent to obtain the constituent length. For stimuli reported to have two constituents, it is possible that the two constituents varied in length.

2.2.3. Constituent size assessment

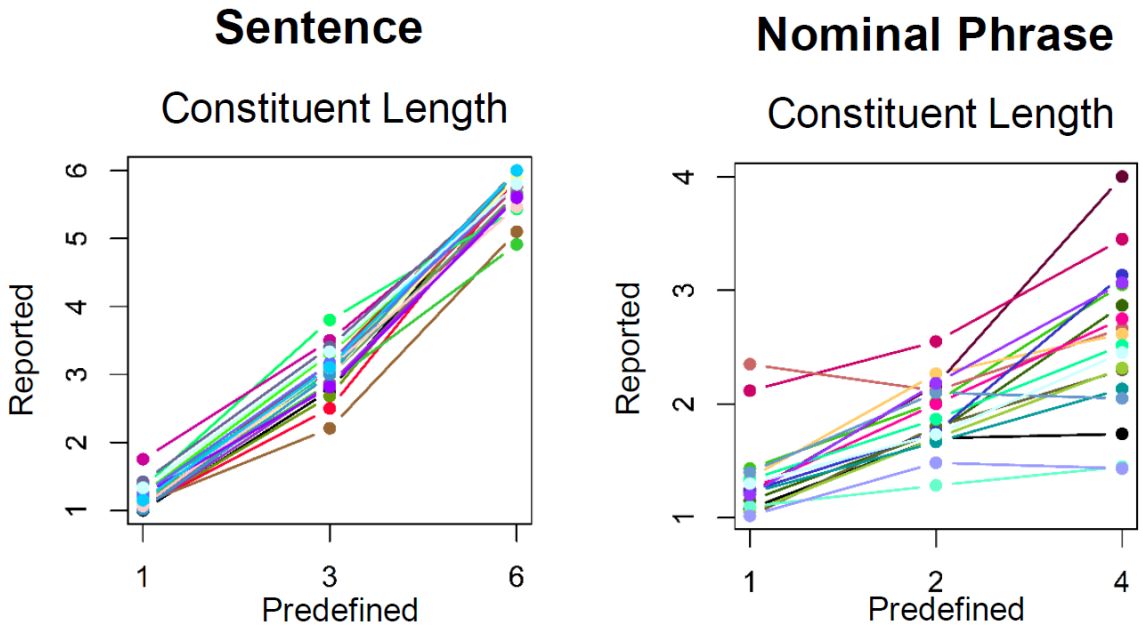


Figure 1. Reported constituent length. The mean reported constituent lengths are shown as a function of the predefined constituent lengths. Each line represents one subject. Reported and predefined constituent lengths were quite consistent for sentences, while larger individual variation was observed for nominal phrases. The reported length was used for further analyses.

As expected, the reported and predefined constituent lengths were fairly congruent for sentences but less so for nominal phrase (Figure 1). The reported lengths were used for further analyses. Length categories not included in our original design were excluded because they were sparsely reported (Appendix). The nominal phrase condition showed a strong inter-individual variability. Some participants reported long constituents more often than the other participants. We computed the mean reported constituent lengths as an index of individual combination tendency and used it for later fMRI analysis.

For the RT analysis of the post-scan assessment, trials with RT shorter than 200 ms, longer than 5000 ms, or exceeding 3 standard deviations from the mean of the individual subject were excluded. For sentences, paired t-tests showed significant difference (Table 2)($p < .05$, Bonferroni correction) between all constituent lengths (medium vs. short: $t(21) = 2.89$; medium vs. long $t(21) = 6.36$; short vs. long: $t(21) = 3.78$). Medium constituents

yielded significantly longer RTs than both long and short constituents. For nominal phrases, no significant difference between constituent lengths was detected. Individual RTs were incorporated into our fMRI model to make sure that the constituent length effect would not be confounded with difficulty.

Table 2 RT of post-scan constituent assessment.

Sentence				
Constituent Length		No. of Constituent	RT (ms)	95% CI
Short	1	6	752	609-894
Medium	3	2	869	740-999
Long	6	1	580	508-652

Nominal Phrase				
Constituent Length		No. of Constituent	RT (ms)	95% CI
Short	1	4	1157	865-1450
Medium	2	2	1122	870-1374
Long	4	1	973	744-1202

2.3. Procedure

Before entering the scanner, the participants were shown a few examples of the stimuli and instructed about the task: They were asked to respond to a few probe stimuli containing commands to press a button*;* this task was chosen to ensure that participants pay attention to the stimuli. After the acquisition of the anatomical scan, the participants received 3 functional runs for the sentence condition and 2 runs for the nominal phrases condition; the order of condition was counterbalanced across participants (except for the first five participants who received only the sentence condition). All runs contained 10 trials for each of the three levels and 8 additional trials containing commands to press a button. The trials were presented in random order. In each trial, the stimulus was displayed word by word (2 characters at a time), using rapid serial presentations (300 msec per word). The inter-trial interval was 12 seconds, during which the participants had to fixate a cross in the middle of the screen. In total, each

subject received 30 trials of each constituent length under the sentence condition and 20 trials of each constituent length under the nominal phrase condition.

After completing the scanning session, out of the scanner, the participants were presented with all the stimuli previously displayed to them during the experiment and asked to assess the number of constituents.

2.4. Imaging parameters

The acquisition of brain images was performed with a 3T Siemens Tim Trio scanner with a 32-channel phased-array head coil. A 1x1x1 T1-weighted anatomical image was first acquired for each participant. For the functional runs, EPI images were acquired with the following parameters: 35 axial slices (Field of View=216mm) with a spatial resolution of 3x3x3 mm (covering the full brain), Time of Repetition=2sec, Time of Echo=30 msec; Flip Angle=90°. 257 functional volumes were acquired for each nominal phrase session, 269 volumes for each sentence session.

2.5. fMRI Analysis

2.5.1. Preprocessing

The probe trials requesting to press a button during scanning were detected 87% (standard deviation: 9%) of the time, showing that participants were attentive. Data processing was performed with SPM8 (Wellcome Department of Cognitive Neurology, software available at http://www.fil.ion.ucl.ac.uk/spm). Functional volumes were first corrected for slice timing differences and realigned to correct for motion correction. The anatomical and functional EPI images were spatially normalized onto the MNI template using the unified segmentation normalization algorithm of SPM8 with its default parameters. Finally, the functional images were smoothed with an isotropic Gaussian kernel (FWHM=5mm).

2.5.2. Whole-brain analysis

The nominal phrase and sentence conditions were analyzed in separate SPM models based on the reported constituent length from each participant (please see the constituent number assessment section for more details and Supplementary 4 for the

results based on the predefined constituent length). Experimental effects at each voxel were assessed using a multi-session design matrix. The design matrix included a regressor modeling all stimuli, accompanied by three parametric modulators: centered log RT, an indicator variable specifying if the item belonged to the medium condition (1 in this case, 0 else), and another one for the long condition. In addition to the regressors of interest, two additional regressors were included for the probe trials and trials with reported length categories not included in our original design stimuli (the original categories are 1, 3, and 6 for sentence. 1, 2, and 4 for nominal phrase)(Appendix), as well as six regressors estimating the movements. All regressors except the movements regressors were obtained by convolving the standard hemodynamic response with a boxcar function starting at stimulus' onset and lasting 1.2s for the noun-phrase condition and 1.8s for the sentence condition (corresponding to the actual duration of the display). We did not orthogonalize the regressors to each other.

Effect maps for long and medium constituent lengths were obtained by setting contrast with value 1 in the relevant column and 0 elsewhere for each participant. Since short constituent, i.e. word list, served as the reference level in this model, the effect of the long constituent modulator reflected the long vs. short difference. The same is for the medium constituent modulator. The contrast images thus obtained from the individual level models were smoothed with a Gaussian kernel (FWHM=8mm) and input to the group level models (one-way within-subject ANOVA) for the nominal phrase condition (N=15, three participants were excluded because no stimuli were reported as one nominal phrase with four words, i.e., long constituent, by them) and sentence conditions (N = 22), respectively. The AAL toolbox was used to label brain regions (Tzourio-Mazoyer et al., 2002).

For the comparison between sentence and nominal phrase conditions, a group model including only the 15 overlapped participants was built (flexible factorial), with four regressors of interest (long/medium constituent X sentence/nominal phrase) and one regressor for each participant in each condition.

To test the individual combination tendency effect, we built another group-level model (multiple-regression) with intercept and individual tendency (mean reported

constituent length) regressors for long vs. short and medium vs. short contrasts, respectively, using the same input images.

Contrast images of the RT effect were obtained from the individual model by setting contrast with value 1 in the RT regressor and 0 elsewhere and tested with a group level one-sample t-test.

2.5.3. ROI analysis

Seven independently defined a priori regions of interest (ROI) were adopted (Figure 2). Left IFGorb, left IFGtri, TP, aSTS, pSTS, and TPJ ROIs were defined as in Pallier et al. (2011). 14 voxels in the left IFGtri ROI overlapped with the anatomical mask of left IFGoper were excluded, while 67 voxels remained. Left IFGoper ROI was defined based on the meta-analysis of Zaccarella, Schell, & Friederici (2017), which reported a cluster showing sentence > content or function word list effect with peak coordinate at MNI [-60, 14 , 12]. The left IFGoper ROI was defined as the overlap between a sphere of 10 mm radius centered at the reported peak and the anatomical mask of left IFGoper provided by the AAL toolbox. We applied the same ROI masks to all the participants. One concern regarding this approach is that different sub-regions within the masks might be recruited for different conditions at the individual level (Blank & Fedorenko, 2017). For ROIs defined by individual localizers, please see Supplementary 5.

A repeated measure ANOVA was run for each ROI with constituent length as the within-subject factor and corrected for multiple comparisons (7 ROIs) using the FDR method (Benjamini & Hochberg, 1995). Within ROIs showing a significant difference between constituent lengths, post-hoc pairwise t-tests were performed and corrected using the FDR method.

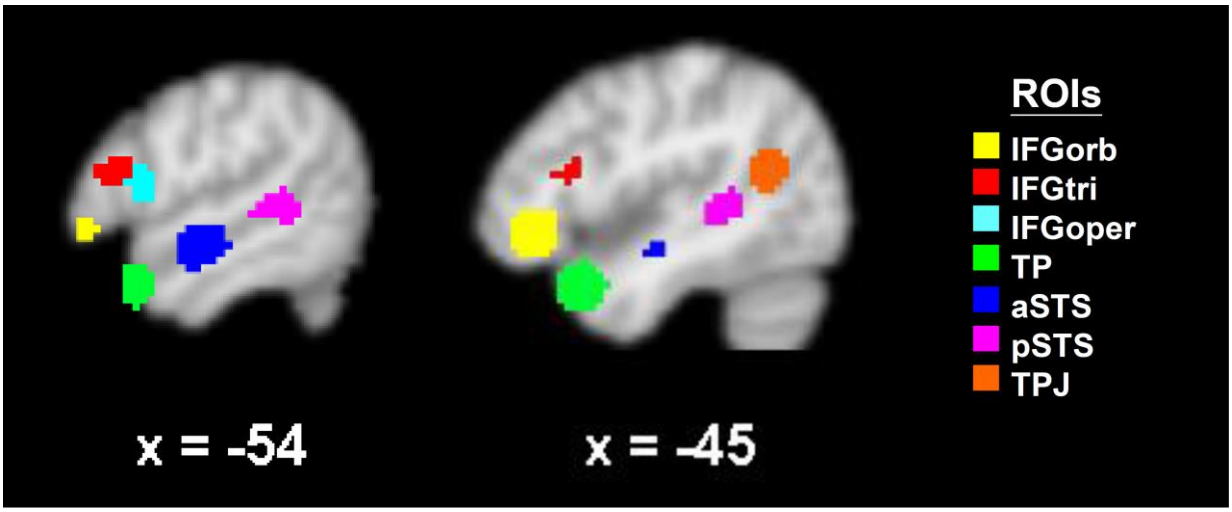


Figure 2. A priori functional ROI masks defined based on previous studies in sentence processing.

3. Results

3.1 Constituent length effect revealed by the whole-brain analysis

First, we compared 6-word sentences (condition “Long”), and two 3-word phrases (condition “Medium”) to lists of 6 unconnected words (condition “Short”) (see Table 1) — both Long vs. Short and Medium vs. Short contrasts markedly overlapped with the findings in French (Pallier et al., 2011) (Figure 3, Table 3, and Figure 2), including the left orbital part of IFG, the triangular part of IFG, the temporal pole, the aSTS, pSTS, and TPJ. However, Medium vs. Short contrast was more extensive and more robust statistically, which was unexpected. Since medium constituents were associated with longer RTs (Table 2), one might wonder whether these findings actually reflected RT differences. However, our fMRI model incorporated RTs as separate regressors and ruled out this RT explanation. The group-level t-test for the RT effect showed no significant results. Please see the discussion section for possible explanations for this finding.

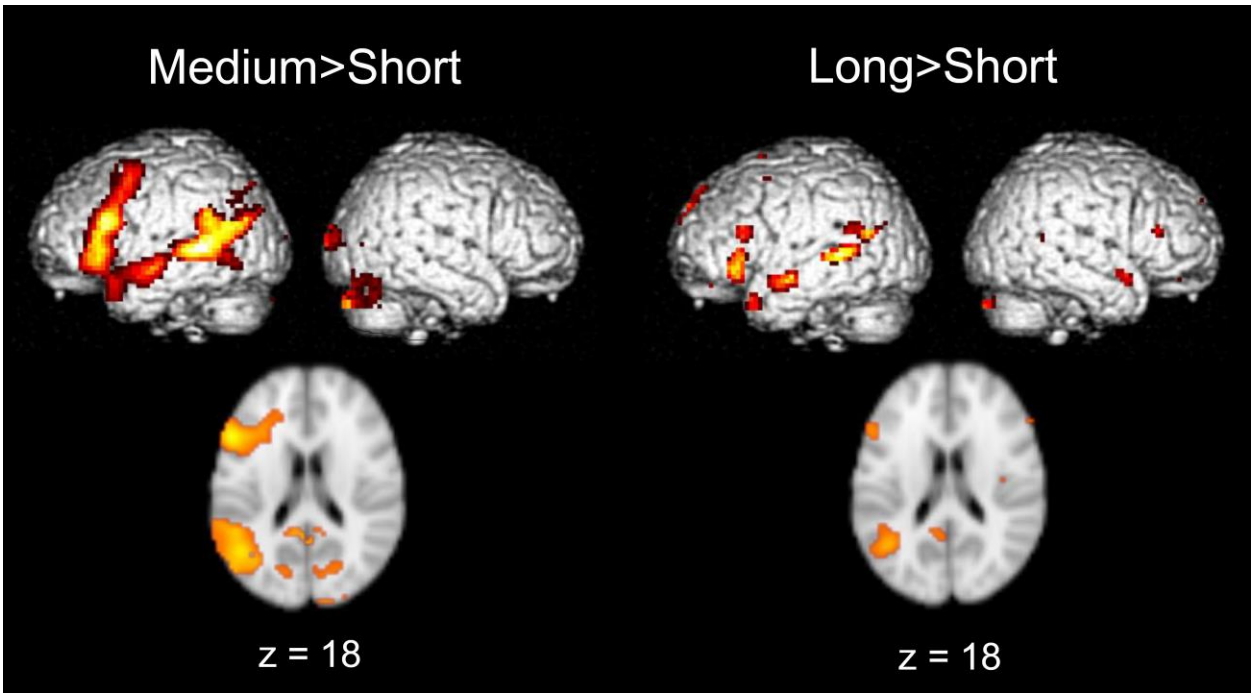


Figure 3. SPMt maps of the contrast between Medium v.s. Short (thresholded at p < .005 voxelwise; cluster size pFWE < .05) and Long v.s. Short constituent (thresholded at p < .005 voxelwise; uncorrected) under sentence condition. Both maps were exclusively masked with negative short constituent effect thresholded at p< .05.

Table 3 Regions showing the effect of constituent length under sentence condition (N=22. Medium v.s. Short: p < .005 voxelwise; cluster size pFWE < .05. Long v.s. Short p < .005 voxelwise; uncorrected. Both were exclusively masked with negative short constituent effect thresholded at p< .05).

Sentence	T-value	x	y	z	Cluster size
Long v.s. Short					
Frontal Inf Orb L	4.68	-39	26	-8	109
Frontal Inf Tri L	3.91	-54	23	19	34
Frontal Sup L	5.46	-12	56	40	13
Frontal Inf Tri R	3.74	57	29	16	11
Temporal Pole Mid L	3.62	-45	14	-32	28
Temporal Pole Mid R	5.48	60	8	-17	22
Temporal Mid L	3.82	-63	-7	-11	87
Temporal Mid L	5.83	-54	-40	4	127
Temporal Mid L	5.11	-42	-58	22	94
Cerebelum Crus2 R	3.55	18	-85	-35	23
Medium v.s. Short					
Frontal Inf Orb L	5.57	-42	26	-5	2557
Frontal Inf Tri L	6.82	-54	20	22	
Frontal Mid L	5.60	-45	8	55	
Temporal Pole Mid L	3.69	-45	14	-26	
Temporal Mid L	5.81	-63	-10	-8	
Temporal Mid L	7.91	-54	-40	4	
Temporal Mid L	4.88	-39	-58	-2	
Temporal Mid L	6.18	-48	-61	16	
Lingual_R	5.67	15	-46	4	1379
Calcarine_L	5.06	-18	-73	13	
Calcarine_R	4.74	21	-73	10	
Precuneus_L	5.29	-6	-52	7	
Cuneus_R	5.44	18	-100	13	
Cerebelum_6_R	5.07	21	-67	-32	
Cerebelum_Crus2_R	5.89	18	-85	-35	

For the nominal phrase, both Long vs. Short and Medium vs. Short contrasts showed an effect in bilateral triangular part of IFG, left pSTS/pSTG, left pMTG, putamen, and the left middle occipital gyrus (Figure 4, Table 4). The effect was stronger in the long

than in the medium condition. The Long vs. Short contrast also detected the bilateral opercular part of IFG, the left superior medial frontal area, the TPJ, the left IPL, the left precentral gyrus, and the right postcentral gyrus. Please see Supplementary 6 for the overlap between Figure 3 and Figure 4.

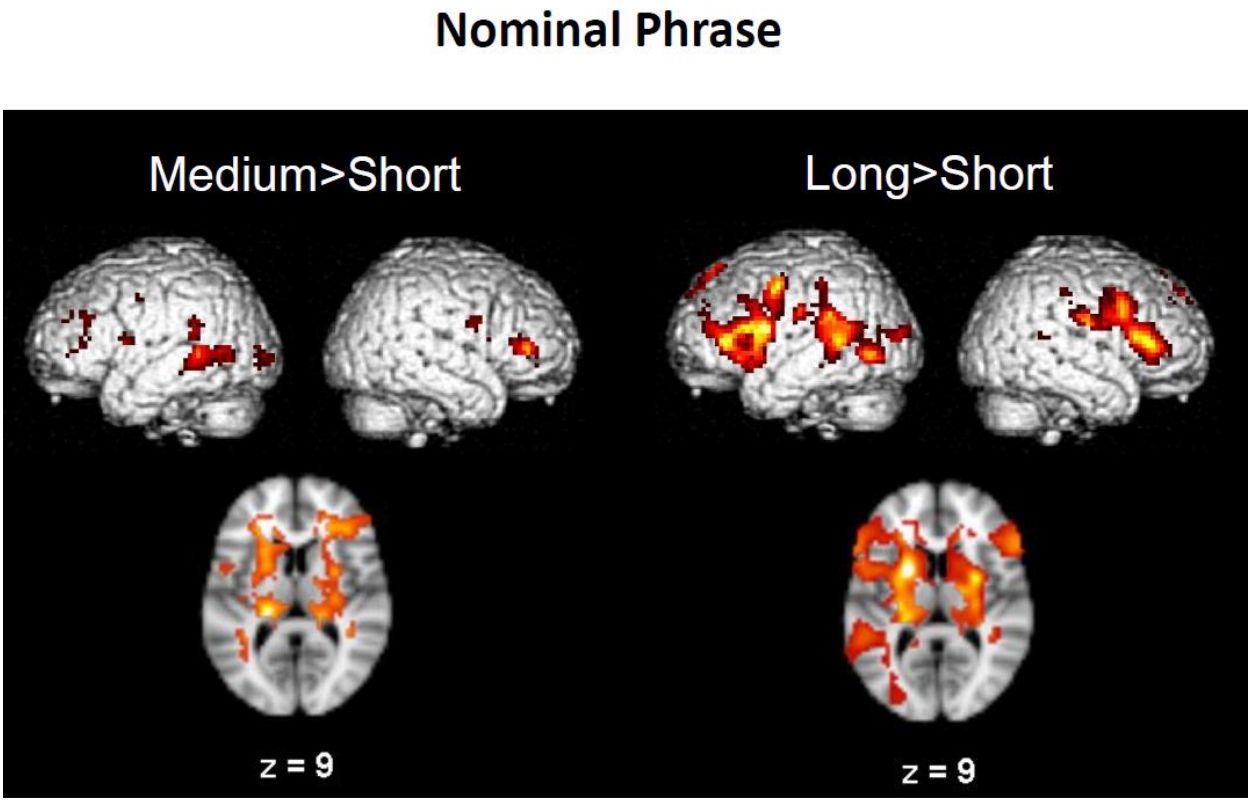
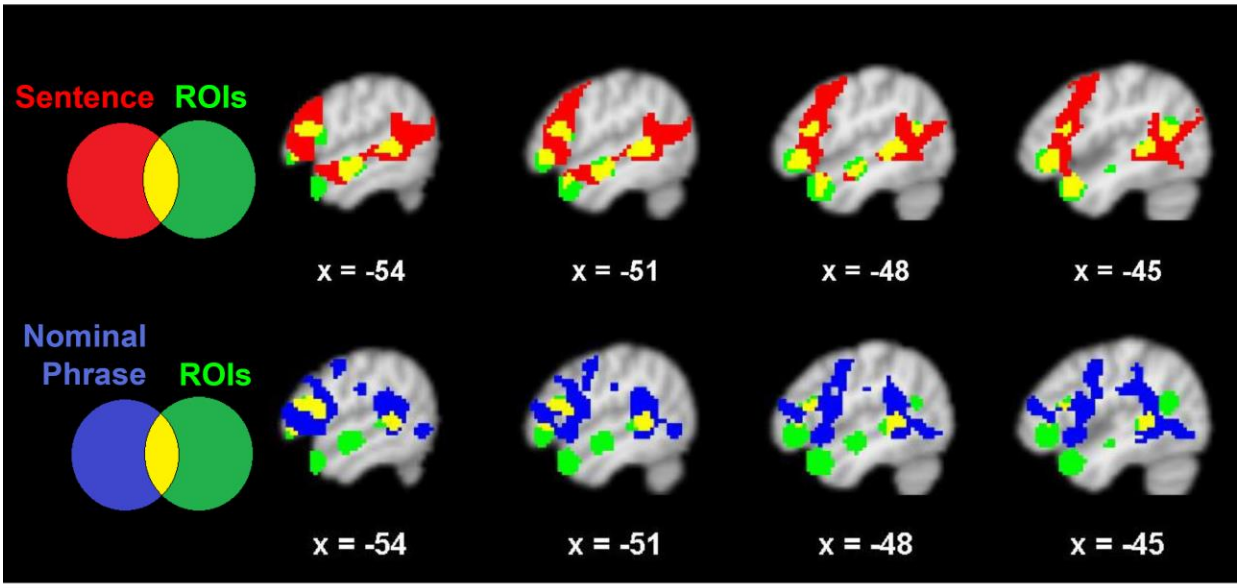


Figure 4. SPMt maps of the contrast between Long v.s. Short and Medium v.s. Short constituent under nominal phrase condition (N=15, thresholded at $p < .005$ voxelwise; cluster size pFWE < .05; exclusively masked with negative short constituent effect thresholded at $p < .05$).

Table 4 Regions showing the effect of constituent length under the nominal phrase condition (N=15, thresholded at p < .005 voxelwise; cluster size pFWE < .05; exclusively masked with negative short constituent effect thresholded at p< .05).

Nominal Phrase	T-value	x	y	z	Cluster size
Long v.s. Short					
Frontal_Inf_Tri_R	6.56	57	23	13	6863
Frontal_Inf_Oper_L	5.61	-57	17	13	
Frontal_Inf_Tri_L	5.45	-57	29	4	
Frontal_Inf_Oper_R	5.35	60	14	31	
Frontal_Sup_Medial_L	4.72	-3	38	58	
Precentral_L	4.87	-51	-4	46	
Postcentral_R	4.29	51	-13	27	
Temporal_Sup_L	5.48	-51	-43	16	
Temporal_Mid_L	4.73	-57	-46	7	
Occipital_Mid_L	3.87	-54	-64	-2	
Putamen_L	11.20	-21	5	13	
Putamen_R	8.36	27	2	13	
Medium v.s. Short					
Frontal Inf Tri R	4.20	54	44	1	3150
Frontal Inf Tri L	3.74	-27	28	7	
Insula R	5.71	36	32	7	
Temporal Mid L	6.90	-45	-40	-5	
Occipital_Mid_L	3.86	-21	85	1	
Putamen L	6.38	-21	5	13	
Thalamus L	6.11	-6	-22	16	

3.2 Constituent length effect in a priori ROIs



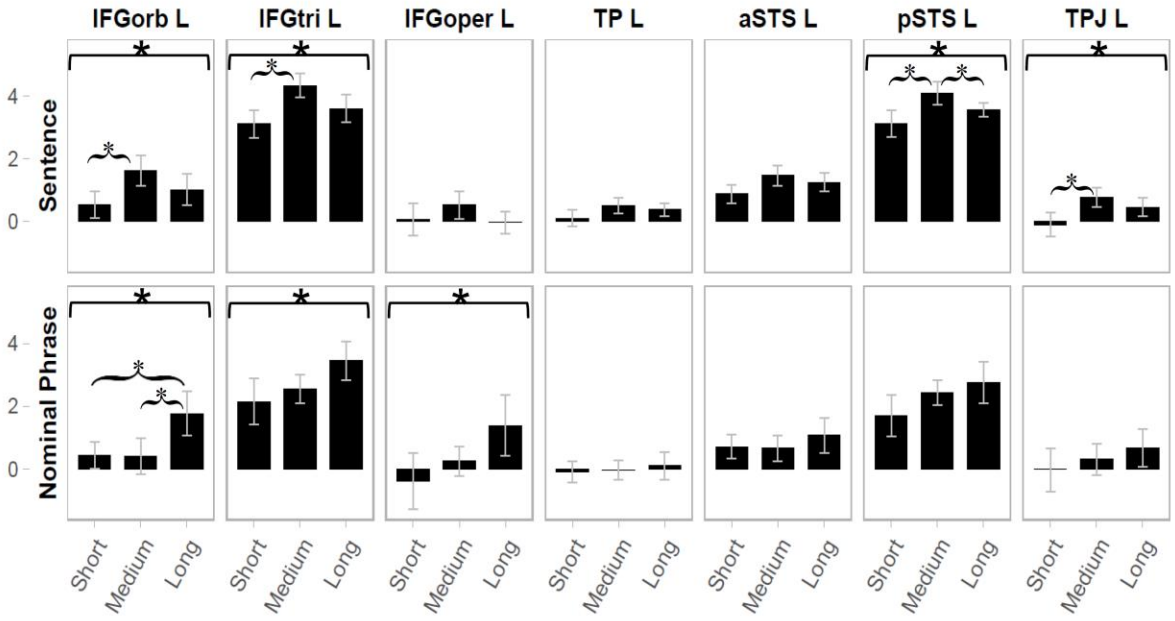


Figure 5. The effect of constituent length in a priori ROIs. Upper: overlap between the ROIs (Figure 2) and the effect of constituent length under sentence and nominal phrase conditions. Lower: the averaged fMRI activations for the three constituent lengths. Error bars indicate 95% confidence interval (CI). Square brackets indicate significant difference between the three constituent lengths ($p < .05$). Curly brackets indicate significant post-hoc pairwise comparisons ($p < 05$, FDR correction).

The effect of constituent length was tested in a priori ROIs defined based on previous studies using sentence stimuli (Figure 5). Under the sentence conditions, difference between constituent lengths was significant in IFGorb ($F(2, 42) = 3.82$, $p = .030$), IFGtri ($F(2, 42)=6.21$, $p = .004$), pSTS ($F(2, 42) = 5.64$, $p = .007$), and TPJ ($F(2, 42) = 5.25$, $p = .009$), but not significant in IFGoper ($F(2,42) = 1.31$, $p = . 280$), aSTS ($F(2, 42) = 2.88$, $p = .067$), and TP ($F(2, 42) = 1.96$, $p = .153$). Among regions showing a significant main effect of constituent size, post-hoc pairwise comparisons were significant (FDR-corrected $p < .05$) in pSTS (medium>short: $t(1,21) = 2.65$, $p = .030$; medium>long: $t(1, 21) = 2.52$, $p = .030$), IFGorb (medium>short: $t(1, 21) = 2.98$, $p = .021$), IFGtri (medium>short: $t(1,21) = 3.63$, $p = .005$), and TPJ (medium>short: $t(1, 21) = 2.94$, $p = .023$).

Under the nominal phrase condition, difference between constituent lengths was significant in IFGorb ($F(2, 28) = 5.56$, $p = .009$) and IFGtri ($F(2, 28) = 3.67$, $p = .038$), and IFGoper ($F(2, 28)=3.80$, $p = .035$), but not significant in pSTS ($F(2, 28) = 2.69$, $p = .086$), TP ($F(2, 28) = .23$, $p = .797$), TPJ ($F(2, 28) = 1.07$, $p = .356$), and aSTS ($F(2, 28) = 0.73$, $p = .490$). Among regions showing a significant main effect of constituent size, post-hoc pairwise comparisons were significant (FDR-corrected $p < .05$) in IFGorb (long>short: $t(1, 14) = 2.85$, $p = .039$; medium>short: $t(1, 14) = 2.40$, $p < .046$).

Consistent with the whole-brain analyses, left IFGorb and IFGtri showed the effect of constituent length under both sentence and nominal phrase conditions. Left pSTS and TPJ only showed the difference between constituent lengths under sentence condition, while left IFGoper showed an effect only under nominal phrase condition.

3.3 Comparison between sentence and nominal phrase

To compare the effects of constituent length under sentence and nominal phrase conditions, we built a group model including only the 15 subjects that participated in both conditions using Long v.s. Short and Medium v.s. Short contrast images from the first-level model. As shown in Figure 6 and Table 5, The construction of nominal phrase showed a stronger effect in bilateral dorsal IFGtri, left IFGoper, left pMTG/pSTG, left IPL,

right superior frontal area, left MFG, left precentral gyrus, putamen, and thalamus, while no area showed significantly larger effect under the sentence condition.

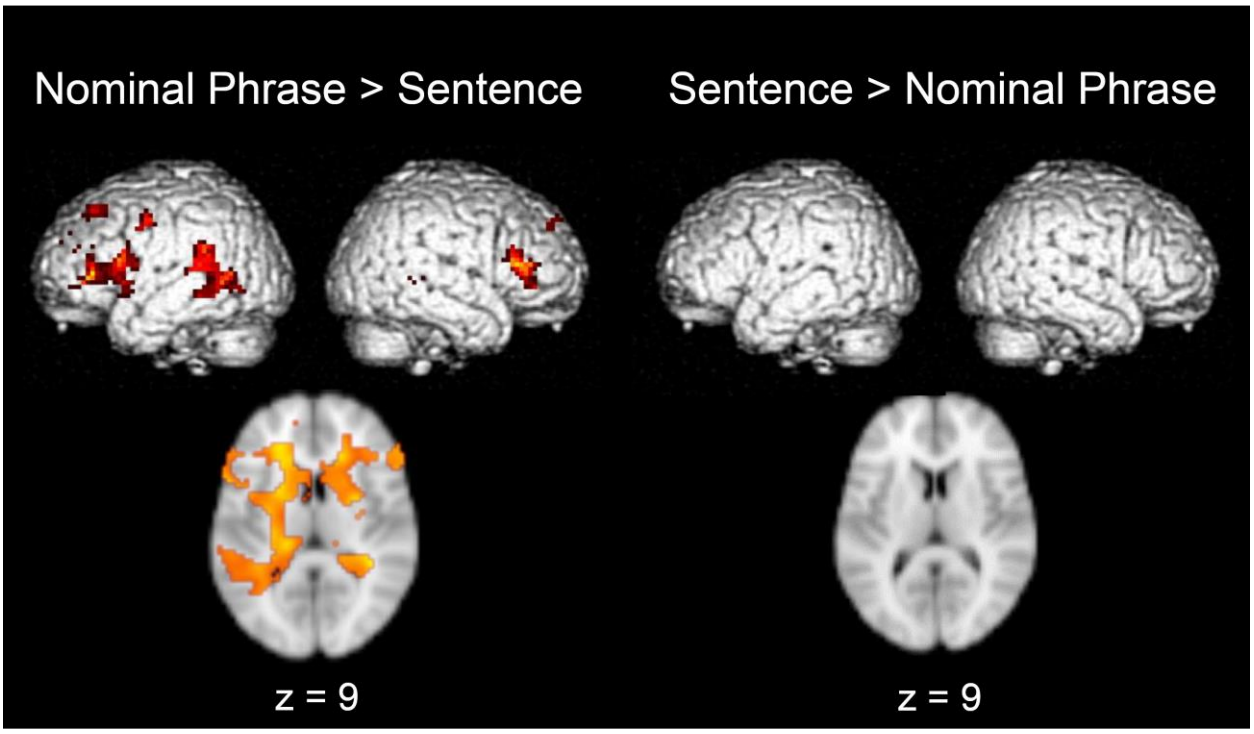


Figure 6. Comparison between the effects of constituent length under sentence and nominal phrase conditions. Left: regions showing larger effect under nominal phrase condition (N=15, thresholded at $p < .005$ voxelwise; cluster size $pFWE < .05$; exclusively masked with negative sentence effect thresholded at $p < .05$). Right: No region showed significantly larger effect under the sentence condition.

Table 5. Regions showing larger effect of constituent length under nominal phrase condition (N=15, thresholded at p < .005 voxelwise; cluster size pFWE < .05; exclusively masked with negative sentence effect thresholded at p< .05).

Nominal Phrase > Sentence	T-value	x	y	z	Cluster size
Frontal_Inf_Oper_L	5.38	-57	14	10	4225
Frontal_Inf_Tri_L	4.13	-57	32	4	
Frontal_Inf_Tri_R	5.06	57	23	16	
Frontal_Mid_L	3.91	-30	32	49	
Frontal_Sup_Medial_R	3.22	9	50	37	
Frontal_Sup_R	4.13	15	38	34	
Precentral_L	4.22	-36	-4	43	
Putamen_L	5.18	-30	5	4	
Rolandic_Oper_L	3.33	-48	8	1	
Temporal_Mid_L	3.97	-45	-58	1	
Temporal_Sup_L	5.21	-45	-43	22	
Thalamus_L	6.19	-18	-28	7	
Thalamus_R	4.42	18	-25	16	

3.4 Individual combination tendency effect

As revealed by the behavioral data (Figure 1), there was more uncertainty about the parsing of nominal phrases, presumably due to the absence of morphosyntactic cues. Regions found only under the nominal phrase condition might be involved to access and integrate other sources of information than syntax, like semantics and world knowledge (Figure 6). Individuals who had better access to non-syntactic information might be better at synthesizing longer phrases. We, therefore, computed the mean reported constituent length for each participant and used it as an index of individual combination tendency. Taking the effect maps of Long vs. Short contrast from individual models as input, a group-level model with intercept and individual combination tendency as regressors were built. The same operation was done for the Medium vs. Short contrast. We detected a positive correlation between brain activity and the behavioral combination tendency effect only in the nominal phrase (Supplementary 7) but not in the sentence condition. No negative correlation was found. A conjunction analysis revealed a combination tendency effect in the upper left triangular and the opercular part of IFG, and in subcortical regions (Figure

7 and Table 6). In other words, participants whose brain activations in these regions increased more with longer constituents were more likely to combine words.

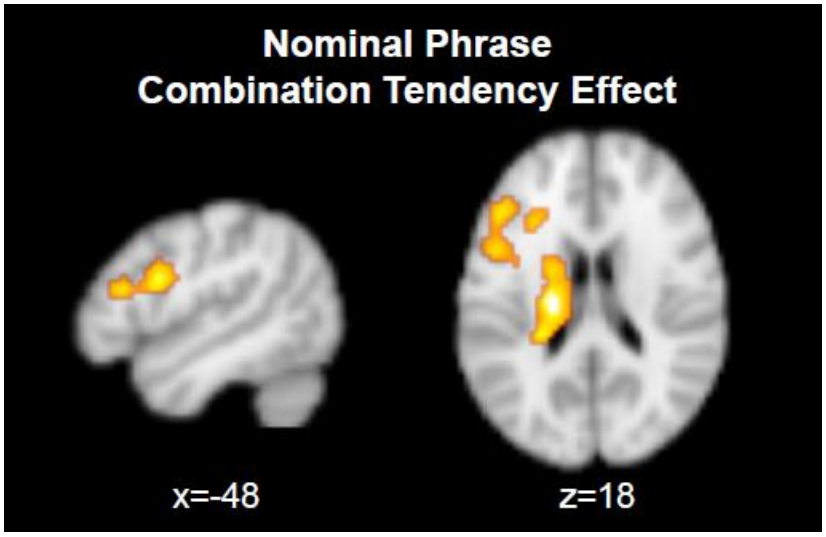


Figure 7. Regions showing individual combination tendency effect under nominal phrase condition for both Medium v.s. Short and Long v.s. Short contrasts revealed by conjunction analysis (N=15; SPMt map thresholded at $p < .005$ voxelwise; cluster size $pFWE < .05$).

Table 6. Regions showing individual combination tendency effect under nominal phrase condition for both Medium v.s. Short and Long v.s. Short contrasts revealed by conjunction analysis (N=15; thresholded at $p < .005$ voxelwise; cluster size $pFWE < .05$).

Nominal Phrase	T-value	x	y	z	Cluster size
Combination tendency					
Frontal_Inf_Oper_L	3.98	-51	8	25	753
Frontal_Inf_Tri_L	3.70	-45	29	16	
Caudate_L	4.73	-21	-16	22	
Thalamus_L	3.13	-6	-22	13	

4. Discussion

Previous studies have consistently shown that compared to word lists, sentences and phrases elicit stronger activation in left IFG and along left STG/STS (Zaccarella et al., 2017), which is attributed to the combinatorial operations involved in constituent building. Combining words into larger units can be facilitated a lot by morphosyntactic cues. This study examined whether the same areas are recruited even when few or no morphosyntactic cues are available. We have two main findings. (1) In Mandarin, we observed that processing sentences yielded stronger activity than word lists in a similar network as languages with rich morphosyntactic cues (Figure 5). (2) Our nominal phrases, which contained no morphosyntactic cues, showed a constituent length effect beyond areas engaged in sentence processing (Figure 6). This might reflect an increased reliance on semantic and pragmatic information.

In the sentence condition, the whole-brain analysis identifies brain regions that remarkably overlap with the network previously reported in other languages (Zaccarella et al., 2017), including the left IFGorb, IFGtri, temporal pole, aSTS, pSTS, and TPJ (Figure 3 and Figure 5). Among them, left IFGorb, IFGtri, pSTS, and TPJ also showed a difference between constituent lengths in the ROI analysis. Such findings indicate that these regions are engaged during constituent encoding across languages regardless of typological differences in the density of morphosyntactic cues.

The nominal phrase condition revealed areas that were not detected in the sentence condition, including the bilateral dorsal IFGtri, left IFGoper, left pMTG/pSTG, left IPL, and several subcortical areas (Figure 6). Nominal phrases were structurally more ambiguous, as reflected by the behavioral results (Figure 1), presumably due to the absence of syntactic constraints on linguistic composition. Under this situation, constituent parsing must depend on non-syntactic information. For example, for the nominal phrase 玫瑰精油身體乳液 (“*rose oil body lotion*”), better access to semantic knowledge (e.g., oil as petroleum, as a cooking ingredient, or as an essential oil), as well as pragmatic, and world knowledge (e.g., the knowledge that only the latter type of oil is used in lotions) could help to build longer constituents. These regions might be engaged in accessing and integrating information from different sources. Studies comparing intact narrative vs. scrambled sentence suggested that these regions were

sensitive to the coherence of the contextual information (Lerner, Honey, Silbert, & Hasson, 2011; Xu, Kemeny, Park, Frattali, & Braun, 2005; Yarkoni, Speer, & Zacks, 2008). A meta-analysis of imaging studies has revealed reliable activations in bilateral IFG and pMTG to sentences containing syntactic or semantic ambiguity (Hagoort & Indefrey, 2014), and to condition with a higher demand on semantic executive control (Noonan, Jefferies, Visser, & Lambon Ralph, 2013). The basal ganglia/thalamus has also been proposed to be involved in solving lexical ambiguity with contextual information (see Chenery, Angwin, & Copland, 2008 and Ketteler, Kastrau, Vohn, & Huber, 2008 for reviews).

The same explanation also applies to the individual combination tendency effect. As shown in Figure 6, nominal phrase activated the left IFG more extensively than sentences. Moreover, participants who showed larger constituent length effects in upper left IFGtri/IFGoper, and left thalamus/caudate tended to build longer constituents (Figure 7). They might benefit more from non-syntactic knowledge through better cognitive control. Individual differences in cognitive control might be more salient when external linguistic inputs provide little constraints on parsing. A similar idea has been proposed in Blank & Fedorenko (2017), which demonstrated a domain-general system that responded to language comprehension task but showed a low inter-subject correlation in responses, which was suggested to reflect a loose tracking of the shared linguistic input. That individual difference in cognitive control affected language comprehension, especially with insufficient syntactic cues, has been widely reported at both behavioral (MacDonald, Just, & Carpenter, 1992) and neural levels (Bornkessel-Schlesewsky, Fiebach, & Friederici, 2004; Novais-Santos et al., 2007).

A universal constituent encoding system shared across languages (Dehaene et al., 2015; Hagoort & Indefrey, 2014), utilizing both morphosyntactic and non-morphosyntactic information, has been proposed. Such a system should show an effect of constituent length under both the sentence and nominal phrase conditions. Using a priori defined ROIs, we found an effect of constituent length under both conditions in left IFGtri and left IFGorb (Figure 5, lower). Such results indicate that these regions can utilize non-morphosyntactic information. Left IFGtri have also been reported in previous studies using jabberwocky stimuli, which had only morphosyntactic cues (Goucha &

Friederici, 2015; Matchin et al., 2017; Pallier et al., 2011), and in study comparing Mandarin minimal phrase containing a function word vs. word list (determiner + classifier vs. classifier + classifier)(Wu, Zaccarella, & Friederici, 2018). In addition, van der Burght et al. (van der Burght, Goucha, Friederici, Kreitewolf, & Hartwigsen, 2019) found higher activation in left IFG when prosodic cues were critical for constituent construction. These findings are in line with the proposed universal constituent encoding system utilizing various sources of information. Alternatively, these findings could suggest a functional subdivision in these areas (Fedorenko, Duncan, & Kanwisher, 2012; Goucha & Friederici, 2015), or partial overlap between functional networks (Fedorenko & Thompson-Schill, 2014; Hagoort, 2014).

An unexpected aspect of the data concerns the stronger brain response to medium than long constituent in the sentence condition (Figure 3 and Figure 5). One possible explanation is a garden path effect. Namely, the reader may have started with a wrong syntactic analysis but had to revise it out later in the sequence. Another explanation is that many medium constituent stimuli contain two verbs/sentences (Supplementary 2), which could also increase the processing difficulty. An example is as follows:

校方　　勉強　　答應　　女孩　選擇　離開

school　reluctantly　concede　girl　choose　leave

“The school reluctantly concedes that the girl...” (possible temporary interpretation)

“The school reluctantly concedes. The girl chooses to leave.” (final interpretation)

Another example is 他們 上門 應徵 教練 示範 解說 (they come apply coach demonstrate explain). One possible temporary interpretation is “*They come to apply for a coaching (position)….*” and the final interpretation is “*They come to apply. The coach demonstrates and explains.*” Both examples could induce garden path effect and have two verbs/sentences. While the increased processing difficulty of medium constituent stimuli is reflected by the longer RT in the assessment task (Table 1), by incorporating RT as a regressor separate from the constituent length, we have ruled out the possibility that this finding only reflected an RT effect. It is possible that before the participants reached the final interpretation, multiple representations were constructed in parallel,

resulting in higher activations for the medium constituent trials (Frisch, Schlesewsky, Saddy, & Alpermann, 2002; Mason, Just, Keller, & Carpenter, 2003). However, the two factors, i.e. garden path effect and the number of verbs/sentences, often covary and our stimuli do not allow a clear dissociation.

It has been argued that a subregion of left IFG, i.e. IFGoper/BA44, is specialized in constituent construction based on morphosyntactic cues (Hagoort & Indefrey, 2014; Zaccarella et al., 2017). Several studies have directly manipulated the availability of morphosyntactic cues in the stimuli. Goucha et al. (Goucha & Friederici, 2015) compared the sentence vs. word-list effects using two types of jabberwocky stimuli, one with derivational morphemes (e.g., *dis-* in the English word *dislike*) and one without. They found BA44 under both conditions and BA45 only when derivational morphemes were preserved. Since derivational morphemes could change the meanings of words, they suggested that BA44 is the core region for constituent construction based on pure morphosyntactic cues (i.e., function words and inflections). Similarly, Shell et al. (2017) compared legal phrases with and without function word (determiner + noun vs. adjective + noun) and found BA44 only in phrases with a function word. However, some previous studies using jabberwocky stimuli, which is supposed to reveal structure building based on morphosyntactic cues, found a strong constituent effect in BA45 but no (Matchin et al., 2017) or weak effect in BA44 (Pallier et al., 2011). Furthermore, compared to the sentence condition, our results showed a stronger effect of constituent length in BA44/IFGoper under the nominal phrase condition (Figure 5 and Figure 6), which contained no morphosyntactic cues. Such results do not support that BA44 is universally specialized in constituent construction purely based on morphosyntactic cues.

Among studies which parametrically manipulated constituent length, constituent length effect was found significant in French (Pallier et al., 2011) and English (Matchin et al., 2017), but it seems weaker or less statistically robust in sign language (Moreno, Limousin, Dehaene, & Pallier, 2018) and Mandarin (Bulut, Hung, Tzeng, & Wu, 2017). The current study provides a potential explanation to the discrepancy between studies, i.e., garden path and individual variation in parsing, in other words, temporary as well as more permanent structural ambiguity. The sparseness of morphosyntactic cues could

increase structural ambiguity and individual differences, which reduce the effectiveness of constituent length manipulation and statistical power. As noted by early researchers in linguistic typology (Greenberg, 1960; Sapir, 1921) and monolinguistic fMRI (Goucha & Friederici, 2015) and ERP studies (Ye, Luo, Friederici, & Zhou, 2006; Yu & Zhang, 2008; Zhang et al., 2013; Zhang, Yu, & Boland, 2010) investigating phrase structure building, the primacy of morphosyntactic cues might be an important dimension along which languages differ from each other. Our study only included Mandarin, so a direct comparison between languages is not feasible. Besides, any two languages always differ in multiple dimensions. Therefore, a more systematic comparison is necessary to examine whether cross-language variation in language processing reflects the difference in the reliance on morphosyntactic cues (Bornkessel-Schlesewsky et al., 2011).

We found that structural ambiguity and individual differences increase with fewer morphosyntactic cues. This phenomenon is not unique to Chinese or nominal phrase. Linguistic ambiguity is proposed to reflect a trade-off between clarity and ease of language production (Piantadosi, Tily, & Gibson, 2012). For example, shorter words are more likely to have a larger number of distinct meanings (polysemy). In other words, by allowing some ambiguity in word meaning, we avoid having to produce long words. Ambiguity could increase both the ease of production (Ferreira, 2008) and learning (Kirby, Tamariz, Cornish, & Smith, 2015). Despite the benefits of ambiguity, it obviously has a cost on comprehension, which can be partially compensated by active prediction and rich context (Christiansen & Chater, 2015; Piantadosi et al., 2012). For example, topic-prominent languages, including Chinese (Huang, 1984; Li & Thompson, 1976, 1981; Liejiong, 2005; Tai, 2013) and most sign languages (Crasborn, Van Der Kooij, And, & De Hoop, 2009; Ingram, 1978; Lillo-Martin, 1986), as well as second-language (Fuller & Gundel, 1987), often explicitly mark the topic of a discourse, which is one of the most important contextual information in language comprehension. In brief, linguistic ambiguity varies between and within languages and future studies need to take this into consideration.

In conclusion, our study showed that the effect of sentences vs. word lists in a brain circuit that largely overlaps with previous studies, including left IFGorb/IFGtri and

areas along the left STS/STG, despite the sparseness of morphosyntactic cues in Mandarin. On the other hand, our nominal phrases involved additional regions, including bilateral dorsal IFGtri, left IFGoper, left pMTG/pSTG, left IPL, and several subcortical areas, which might support the access and integration of semantic and pragmatic information in the absence of syntactic constraints on linguistic composition. Our findings emphasize the contributions of both syntactic and non-syntactic information in language processing. New studies will be needed to clarify how their weightings vary with their availabilities and individual capacities.

Acknowledgments
This work was partly supported by the Ministry of Science and Technology, Taiwan [grant number 102-2911-I-010 -507 -, 103-2911-I-010 -503 -, and 104-2410-H-010 -003 -MY2] and the French Agence Nationale de la Rercherche (ANR CONSTRUCT 140301) We gratefully acknowledge support by INSERM, CEA, Collège de France, and the NeuroSpin staff.

References

Bates, E., & Macwhinney, B. (1989). Functionalism and the Competition Model. *Driven & Fried British Functionalism*, 3–76.

Benjamini, Y., & Hochberg, Y. (1995). Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society. Series B (Methodological)*, *57*(1). https://doi.org/10.2307/2346101

Blank, I. A., & Fedorenko, E. (2017). Domain-General Brain Regions Do Not Track Linguistic Input as Closely as Language-Selective Regions. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *37*(41), 9999–10011. https://doi.org/10.1523/JNEUROSCI.3642-16.2017

Bornkessel-Schlesewsky, I., Fiebach, C. J., & Friederici, A. D. (2004). On the cost of syntactic ambiguity in human language comprehension: An individual differences approach. *Cognitive Brain Research*, *21*(1), 11–21. https://doi.org/10.1016/j.cogbrainres.2004.05.007

Bornkessel-Schlesewsky, I., Kretzschmar, F., Tune, S., Wang, L., Genç, S., Philipp, M., … Schlesewsky, M. (2011). Think globally: Cross-linguistic variation in electrophysiological activity during sentence comprehension. *Brain and Language*, *117*(3), 133–152. https://doi.org/10.1016/j.bandl.2010.09.010

Bulut, T., Hung, Y. H., Tzeng, O. J. L., & Wu, D. H. (2017). Neural correlates of processing sentences and compound words in Chinese. *PLoS ONE*, *12*(12), 1–18. https://doi.org/10.1371/journal.pone.0188526

Chenery, H. J., Angwin, A. J., & Copland, D. A. (2008). The basal ganglia circuits, dopamine, and ambiguous word processing: A neurobiological account of priming studies in Parkinson's disease. *Journal of the International Neuropsychological Society*, *14*(3), 351–364. https://doi.org/10.1017/S1355617708080491

Christiansen, M. H., & Chater, N. (2015). The Now-or-Never bottleneck: A fundamental constraint on language. *Behavioral and Brain Sciences*, *39*. https://doi.org/10.1017/S0140525X1500031X

Crasborn, O., Van Der Kooij, E., And, J. R., & De Hoop, H. (2009). Topic agreement in NGT (Sign Language of the Netherlands) 1. *The Linguistic Review*, *26*, 355–370. https://doi.org/10.1515/tlir.2009.013

Dehaene, S., Meyniel, F., Wacongne, C., Wang, L., & Pallier, C. (2015). The Neural Representation of Sequences: From Transition Probabilities to Algebraic Patterns and Linguistic Trees. *Neuron*, *88*(1), 2–19. https://doi.org/10.1016/j.neuron.2015.09.019

Ding, N., Melloni, L., Zhang, H., Tian, X., & Poeppel, D. (2015). Cortical tracking of hierarchical linguistic structures in connected speech. *Nature Neuroscience*, *19*(1), 158–164. https://doi.org/10.1038/nn.4186

Ellis, N. (2002). Frequency effects in language processing. *Studies in Second Language Acquisition*, 143–188.

Fedorenko, E., Duncan, J., & Kanwisher, N. (2012). Language-selective and domain-general regions lie side by side within Broca's area. *Current Biology*, *22*(21), 2059–2062. https://doi.org/10.1016/j.cub.2012.09.011

Fedorenko, E., & Thompson-Schill, S. L. (2014, March 1). Reworking the language network. *Trends in Cognitive Sciences*. Elsevier Current Trends. https://doi.org/10.1016/j.tics.2013.12.006

Ferreira, V. S. (2008). Ambiguity, Accessibility, and a Division of Labor for Communicative Success. *Learning and Motivation*, *49*, 209–246. https://doi.org/10.1016/S0079-7421(08)00006-6

Frisch, S., Schlesewsky, M., Saddy, D., & Alpermann, A. (2002). The P600 as an indicator of syntactic ambiguity. *Cognition*, *85*(3), B83–B92. https://doi.org/10.1016/S0010-0277(02)00126-9

Fuller, J. W., & Gundel, J. K. (1987). Topic-Prominence in Interlanguage. *Language Learning*, *37*(1), 1–18. https://doi.org/10.1111/j.1467-1770.1968.tb01310.x

Gibson, E., & Pearlmutter, N. J. (1998). Constraints on sentence comprehension. *Trends in Cognitive Sciences*, *2*(7), 262–268. https://doi.org/10.1016/S1364-6613(98)01187-5

Goucha, T., & Friederici, A. D. (2015). The language skeleton after dissecting meaning: A functional segregation within Broca's Area. *NeuroImage*, *114*, 294–302. https://doi.org/10.1016/j.neuroimage.2015.04.011

Greenberg, J. H. (1960). A Quantitative Approach to the Morphological Typology of Language. *Source: International Journal of American Linguistics*, *26*(3), 178–194.

https://doi.org/10.1086/464575

Hagoort, P. (2005). On Broca, brain, and binding: A new framework. *Trends in Cognitive Sciences*, *9*(9), 416–423. https://doi.org/10.1016/j.tics.2005.07.004

Hagoort, P. (2014). Nodes and networks in the neural architecture for language: Broca's region and beyond. *Current Opinion in Neurobiology*, *28*, 136–141. https://doi.org/10.1016/J.CONB.2014.07.013

Hagoort, P., & Indefrey, P. (2014). The Neurobiology of Language Beyond Single Words. *Annual Review of Neuroscience*, *37*(1), 347–362. https://doi.org/10.1146/annurev-neuro-071013-013847

Huang, C.-T. J. (1984). On the Distribution and References of Empty Pronouns. In *Linguistics Inquiry* (Vol. 15, pp. 531–574).

Ingram, R. M. (1978). Theme, Rheme, Topic and Comment in the Syntax of American Sign Language. *Sign Language Studies*, *20*, 193–218. https://doi.org/10.1353/sls.1978.0006

Jackendoff, R. (1999). The representational structures of the language faculty and their interactions. In *The neurocognition of language* (pp. 37–79). https://doi.org/10.1093/acprof:oso/9780198507932.003.0003

Ketteler, D., Kastrau, F., Vohn, R., & Huber, W. (2008). The subcortical role of language processing. High level linguistic features such as ambiguity-resolution and the human brain; an fMRI study. *NeuroImage*, *39*(4), 2002–2009. https://doi.org/10.1016/j.neuroimage.2007.10.023

Kirby, S., Tamariz, M., Cornish, H., & Smith, K. (2015). Compression and communication in the cultural evolution of linguistic structure. *Cognition*, *141*, 87–102. https://doi.org/10.1016/j.cognition.2015.03.016

Kuperberg, G. R. (2007). Neural mechanisms of language comprehension: Challenges to syntax. *Brain Research*, *1146*(1), 23–49. https://doi.org/10.1016/j.brainres.2006.12.063

Lerner, Y., Honey, C. J., Silbert, L. J., & Hasson, U. (2011). Topographic mapping of a hierarchy of temporal receptive windows using a narrated story. *Journal of Neuroscience*, *31*(8), 2906–2915. https://doi.org/10.1523/JNEUROSCI.3684-10.2011

Li, C. N., & Thompson, S. a. (1976). Subject and Topic: a New Typology of Language. In C. N. Li (Ed.), *Subject and topic* (pp. 457–489). New York: Academic Press.

Li, C. N., & Thompson, S. A. (1981). *Mandarin: A Functional Reference Grammar* (Vol. null).

Liejiong, X. (2005). Topicalization in Asian Languages. *The Blackwell Companion to Syntax*, *I*.

Lillo-Martin, D. (1986). Two kinds of null arguments in American Sign Language. *Natural Language and Linguistic Theory*, *4*(4), 415–444. https://doi.org/10.1007/BF00134469

MacDonald, M. C., Just, M. A., & Carpenter, P. A. (1992). Working memory constraints on the processing of syntactic ambiguity. *Cognitive Psychology*, *24*(1), 56–98. https://doi.org/10.1016/0010-0285(92)90003-K

Martin, A. E. (2016). Language Processing as Cue Integration: Grounding the Psychology of Language in Perception and Neurophysiology. *Frontiers in Psychology*, *7*, 120. https://doi.org/10.3389/fpsyg.2016.00120

Mason, R. A., Just, M. A., Keller, T. A., & Carpenter, P. A. (2003). Ambiguity in the Brain: What Brain Imaging Reveals About the Processing of Syntactically Ambiguous Sentences. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *29*(6), 1319–1338. https://doi.org/10.1037/0278-7393.29.6.1319

Matchin, W., Hammerly, C., & Lau, E. (2017). The role of the IFG and pSTS in syntactic prediction: Evidence from a parametric study of hierarchical structure in fMRI. *Cortex*, *88*, 106–123. https://doi.org/10.1016/j.cortex.2016.12.010

Matisoff, J. (1991). Sino-Tibetan Linguistics: Present State and Future Prospects. *Annual Review of Anthropology*, *20*, 469–504.

Moreno, A., Limousin, F., Dehaene, S., & Pallier, C. (2018). Brain correlates of constituent structure in sign language comprehension. *NeuroImage*, *167*(April 2017), 151–161. https://doi.org/10.1016/j.neuroimage.2017.11.040

Nelson, M. J., El Karoui, I., Giber, K., Yang, X., Cohen, L., Koopman, H., … Dehaene, S. (2017). Neurophysiological dynamics of phrase-structure building during sentence processing. *Proceedings of the National Academy of Sciences*, *114*(18), E3669–E3678. https://doi.org/10.1073/pnas.1701590114

Noonan, K. A., Jefferies, E., Visser, M., & Lambon Ralph, M. A. (2013). Going beyond inferior prefrontal involvement in semantic control: evidence for the additional contribution of dorsal angular gyrus and posterior middle temporal cortex. *Journal of Cognitive Neuroscience*, *25*(11), 1824–1850. https://doi.org/10.1162/jocn_a_00442

Novais-Santos, S., Gee, J., Shah, M., Troiani, V., Work, M., & Grossman, M. (2007). Resolving sentence ambiguity with planning and working memory resources: Evidence from fMRI. *NeuroImage*, *37*(1), 361–378. https://doi.org/10.1016/j.neuroimage.2007.03.077

Pallier, C., Devauchelle, A.-D., & Dehaene, S. (2011). Cortical representation of the constituent structure of sentences. *Proceedings of the National Academy of Sciences*, *108*(6), 2522–2527. https://doi.org/10.1073/pnas.1018711108

Piantadosi, S. T., Tily, H., & Gibson, E. (2012). The communicative function of ambiguity in language. *Cognition*, *122*(3), 280–291. https://doi.org/10.1016/j.cognition.2011.10.004

Sapir, E. (1921). *Language. An Introduction to the Study of Speech*. New York: Harcourt, Brace & Co. https://doi.org/10.1016/j.livsci.2008.10.015

Schell, M., Zaccarella, E., & Friederici, A. D. (2017). Differential cortical contribution of syntax and semantics: An fMRI study on two-word phrasal processing. *Cortex*, *96*, 105–120. https://doi.org/10.1016/J.CORTEX.2017.09.002

Seidenberg, M., & MacDonald, M. (1999). A probabilistic constraints approach to language acquisition and processing. *Cognitive Science*, *23*(4), 569–588.

Tai, J. H.-Y. (2005). Conceptual Structure and Conceptualizations in Chinese Grammar. *Language and Linguistics*, *6*(4), 539–574.

Tai, J. H.-Y. (2013). Reflections on Typological Characterization of Chinese Grammar. *Human Language Resources and Linguistic Typology*, 59–88.

Thurgood, G., & LaPolla, R. J. (2003). *The Sino-Tibetan languages*. Routledge.

Townsend, D. J., Bever, T. G., & Crocker, M. W. (2001). *Sentence Comprehension : The Integration of Habits and Rules*. *Computational Linguistics* (Vol. 28).

van der Burght, C. L., Goucha, T., Friederici, A. D., Kreitewolf, J., & Hartwigsen, G. (2019). Intonation guides sentence processing in the left inferior frontal gyrus.

Cortex, *117*, 122–134. https://doi.org/10.1016/j.cortex.2019.02.011

Wu, C.-Y., Zaccarella, E., & Friederici, A. D. (2018). Universal neural basis of structure building evidenced by network modulations emerging from Broca's area: The case of Chinese. *Human Brain Mapping*. https://doi.org/10.1002/hbm.24482

Xu, J., Kemeny, S., Park, G., Frattali, C., & Braun, A. (2005). Language in context: Emergent features of word, sentence, and narrative comprehension. *NeuroImage*, *25*(3), 1002–1015. https://doi.org/10.1016/j.neuroimage.2004.12.013

Yarkoni, T., Speer, N. K., & Zacks, J. M. (2008). Neural substrates of narrative comprehension and memory. *NeuroImage*, *41*(4), 1408–1425. https://doi.org/10.1016/J.NEUROIMAGE.2008.03.062

Ye, Z., Luo, Y., Friederici, A. D., & Zhou, X. (2006). Semantic and syntactic processing in Chinese sentence comprehension: Evidence from event-related potentials. *Brain Research*, *1071*(1), 186–196. https://doi.org/10.1016/j.brainres.2005.11.085

Yu, J., & Zhang, Y. (2008). When Chinese semantics meets failed syntax. *Neuroreport*, *19*(7), 745–749. https://doi.org/10.1097/WNR.0b013e3282fda21d

Zaccarella, E., Schell, M., & Friederici, A. D. (2017). Reviewing the functional basis of the syntactic Merge mechanism for language: A coordinate-based activation likelihood estimation meta-analysis. *Neuroscience & Biobehavioral Reviews*, *80*, 646–656. https://doi.org/10.1016/J.NEUBIOREV.2017.06.011

Zhang, Y., Li, P., Piao, Q., Liu, Y., Huang, Y., & Shu, H. (2013). Syntax does not necessarily precede semantics in sentence processing: ERP evidence from Chinese. *Brain and Language*, *126*(1), 8–19. https://doi.org/10.1016/j.bandl.2013.04.001

Zhang, Y., Yu, J., & Boland, J. E. (2010). Semantics does not need a processing license from syntax in reading Chinese. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *36*(3), 765–781. https://doi.org/10.1037/a0019254

Zhou, X., Ye, Z., Cheung, H., & Chen, H.-C. (2009). Processing the Chinese language: An introduction. *Language and Cognitive Processes*, *24*(7–8), 929–946. https://doi.org/10.1080/01690960903201281

Benjamini, Y., & Hochberg, Y. (1995). Controlling the False Discovery Rate: A Practical

and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society. Series B (Methodological)*, *57*(1). https://doi.org/10.2307/2346101

Blank, I. A., & Fedorenko, E. (2017). Domain-General Brain Regions Do Not Track Linguistic Input as Closely as Language-Selective Regions. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *37*(41), 9999–10011. https://doi.org/10.1523/JNEUROSCI.3642-16.2017

Bornkessel-Schlesewsky, I., Fiebach, C. J., & Friederici, A. D. (2004). On the cost of syntactic ambiguity in human language comprehension: An individual differences approach. *Cognitive Brain Research*, *21*(1), 11–21. https://doi.org/10.1016/j.cogbrainres.2004.05.007

Bornkessel-Schlesewsky, I., Kretzschmar, F., Tune, S., Wang, L., Genç, S., Philipp, M., … Schlesewsky, M. (2011). Think globally: Cross-linguistic variation in electrophysiological activity during sentence comprehension. *Brain and Language*, *117*(3), 133–152. https://doi.org/10.1016/j.bandl.2010.09.010

Bulut, T., Hung, Y. H., Tzeng, O. J. L., & Wu, D. H. (2017). Neural correlates of processing sentences and compound words in Chinese. *PLoS ONE*, *12*(12), 1–18. https://doi.org/10.1371/journal.pone.0188526

Chenery, H. J., Angwin, A. J., & Copland, D. A. (2008). The basal ganglia circuits, dopamine, and ambiguous word processing: A neurobiological account of priming studies in Parkinson's disease. *Journal of the International Neuropsychological Society*, *14*(3), 351–364. https://doi.org/10.1017/S1355617708080491

Christiansen, M. H., & Chater, N. (2015). The Now-or-Never bottleneck: A fundamental constraint on language. *Behavioral and Brain Sciences*, *39*. https://doi.org/10.1017/S0140525X1500031X

Crasborn, O., Van Der Kooij, E., And, J. R., & De Hoop, H. (2009). Topic agreement in NGT (Sign Language of the Netherlands) 1. *The Linguistic Review*, *26*, 355–370. https://doi.org/10.1515/tlir.2009.013

Dehaene, S., Meyniel, F., Wacongne, C., Wang, L., & Pallier, C. (2015). The Neural Representation of Sequences: From Transition Probabilities to Algebraic Patterns and Linguistic Trees. *Neuron*, *88*(1), 2–19. https://doi.org/10.1016/j.neuron.2015.09.019

Ding, N., Melloni, L., Zhang, H., Tian, X., & Poeppel, D. (2015). Cortical tracking of hierarchical linguistic structures in connected speech. *Nature Neuroscience*, *19*(1), 158–164. https://doi.org/10.1038/nn.4186

Ellis, N. (2002). Frequency effects in language processing. *Studies in Second Language Acquisition*, 143–188.

Fedorenko, E., Duncan, J., & Kanwisher, N. (2012). Language-selective and domain-general regions lie side by side within Broca's area. *Current Biology*, *22*(21), 2059–2062. https://doi.org/10.1016/j.cub.2012.09.011

Fedorenko, E., & Thompson-Schill, S. L. (2014, March 1). Reworking the language network. *Trends in Cognitive Sciences*. Elsevier Current Trends. https://doi.org/10.1016/j.tics.2013.12.006

Ferreira, V. S. (2008). Ambiguity, Accessibility, and a Division of Labor for Communicative Success. *Learning and Motivation*, *49*, 209–246. https://doi.org/10.1016/S0079-7421(08)00006-6

Frisch, S., Schlesewsky, M., Saddy, D., & Alpermann, A. (2002). The P600 as an indicator of syntactic ambiguity. *Cognition*, *85*(3), B83–B92. https://doi.org/10.1016/S0010-0277(02)00126-9

Fuller, J. W., & Gundel, J. K. (1987). Topic-Prominence in Interlanguage. *Language Learning*, *37*(1), 1–18. https://doi.org/10.1111/j.1467-1770.1968.tb01310.x

Gibson, E., & Pearlmutter, N. J. (1998). Constraints on sentence comprehension. *Trends in Cognitive Sciences*, *2*(7), 262–268. https://doi.org/10.1016/S1364-6613(98)01187-5

Goucha, T., & Friederici, A. D. (2015). The language skeleton after dissecting meaning: A functional segregation within Broca's Area. *NeuroImage*, *114*, 294–302. https://doi.org/10.1016/j.neuroimage.2015.04.011

Greenberg, J. H. (1960). A Quantitative Approach to the Morphological Typology of Language. *Source: International Journal of American Linguistics*, *26*(3), 178–194. https://doi.org/10.1086/464575

Hagoort, P. (2005). On Broca, brain, and binding: A new framework. *Trends in Cognitive Sciences*, *9*(9), 416–423. https://doi.org/10.1016/j.tics.2005.07.004

Hagoort, P. (2014). Nodes and networks in the neural architecture for language: Broca's

region and beyond. *Current Opinion in Neurobiology*, *28*, 136–141. https://doi.org/10.1016/J.CONB.2014.07.013

Hagoort, P., & Indefrey, P. (2014). The Neurobiology of Language Beyond Single Words. *Annual Review of Neuroscience*, *37*(1), 347–362. https://doi.org/10.1146/annurev-neuro-071013-013847

Huang, C.-T. J. (1984). On the Distribution and References of Empty Pronouns. In *Linguistics Inquiry* (Vol. 15, pp. 531–574).

Ingram, R. M. (1978). Theme, Rheme, Topic and Comment in the Syntax of American Sign Language. *Sign Language Studies*, *20*, 193–218. https://doi.org/10.1353/sls.1978.0006

Jackendoff, R. (1999). The representational structures of the language faculty and their interactions. In *The neurocognition of language* (pp. 37–79). https://doi.org/10.1093/acprof:oso/9780198507932.003.0003

Ketteler, D., Kastrau, F., Vohn, R., & Huber, W. (2008). The subcortical role of language processing. High level linguistic features such as ambiguity-resolution and the human brain; an fMRI study. *NeuroImage*, *39*(4), 2002–2009. https://doi.org/10.1016/j.neuroimage.2007.10.023

Kirby, S., Tamariz, M., Cornish, H., & Smith, K. (2015). Compression and communication in the cultural evolution of linguistic structure. *Cognition*, *141*, 87–102. https://doi.org/10.1016/j.cognition.2015.03.016

Kriegeskorte, N., Simmons, W. K., Bellgowan, P. S. F., & Baker, C. I. (2009). Circular analysis in systems neuroscience: the dangers of double dipping. *Nature Neuroscience*, *12*(5), 535–540. https://doi.org/10.1038/nn.2303

Kuperberg, G. R. (2007). Neural mechanisms of language comprehension: Challenges to syntax. *Brain Research*, *1146*(1), 23–49. https://doi.org/10.1016/j.brainres.2006.12.063

Li, C. N., & Thompson, S. a. (1976). Subject and Topic: a New Typology of Language. In C. N. Li (Ed.), *Subject and topic* (pp. 457–489). New York: Academic Press.

Li, C. N., & Thompson, S. A. (1981). *Mandarin: A Functional Reference Grammar* (Vol. null).

Liejiong, X. (2005). Topicalization in Asian Languages. *The Blackwell Companion to*

Syntax, *I*.

Lillo-Martin, D. (1986). Two kinds of null arguments in American Sign Language. *Natural Language and Linguistic Theory*, *4*(4), 415–444. https://doi.org/10.1007/BF00134469

MacDonald, M. C., Just, M. A., & Carpenter, P. A. (1992). Working memory constraints on the processing of syntactic ambiguity. *Cognitive Psychology*, *24*(1), 56–98. https://doi.org/10.1016/0010-0285(92)90003-K

Martin, A. E. (2016). Language Processing as Cue Integration: Grounding the Psychology of Language in Perception and Neurophysiology. *Frontiers in Psychology*, *7*, 120. https://doi.org/10.3389/fpsyg.2016.00120

Mason, R. A., Just, M. A., Keller, T. A., & Carpenter, P. A. (2003). Ambiguity in the Brain: What Brain Imaging Reveals About the Processing of Syntactically Ambiguous Sentences. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *29*(6), 1319–1338. https://doi.org/10.1037/0278-7393.29.6.1319

Matchin, W., Hammerly, C., & Lau, E. (2017). The role of the IFG and pSTS in syntactic prediction: Evidence from a parametric study of hierarchical structure in fMRI. *Cortex*, *88*, 106–123. https://doi.org/10.1016/j.cortex.2016.12.010

Matisoff, J. (1991). Sino-Tibetan Linguistics: Present State and Future Prospects. *Annual Review of Anthropology*, *20*, 469–504.

Moreno, A., Limousin, F., Dehaene, S., & Pallier, C. (2018). Brain correlates of constituent structure in sign language comprehension. *NeuroImage*, *167*(April 2017), 151–161. https://doi.org/10.1016/j.neuroimage.2017.11.040

Nelson, M. J., El Karoui, I., Giber, K., Yang, X., Cohen, L., Koopman, H., … Dehaene, S. (2017). Neurophysiological dynamics of phrase-structure building during sentence processing. *Proceedings of the National Academy of Sciences*, *114*(18), E3669–E3678. https://doi.org/10.1073/pnas.1701590114

Noonan, K. A., Jefferies, E., Visser, M., & Lambon Ralph, M. A. (2013). Going beyond inferior prefrontal involvement in semantic control: evidence for the additional contribution of dorsal angular gyrus and posterior middle temporal cortex. *Journal of Cognitive Neuroscience*, *25*(11), 1824–1850. https://doi.org/10.1162/jocn_a_00442

Novais-Santos, S., Gee, J., Shah, M., Troiani, V., Work, M., & Grossman, M. (2007). Resolving sentence ambiguity with planning and working memory resources: Evidence from fMRI. *NeuroImage*, *37*(1), 361–378. https://doi.org/10.1016/j.neuroimage.2007.03.077

Pallier, C., Devauchelle, A.-D., & Dehaene, S. (2011). Cortical representation of the constituent structure of sentences. *Proceedings of the National Academy of Sciences*, *108*(6), 2522–2527. https://doi.org/10.1073/pnas.1018711108

Piantadosi, S. T., Tily, H., & Gibson, E. (2012). The communicative function of ambiguity in language. *Cognition*, *122*(3), 280–291. https://doi.org/10.1016/j.cognition.2011.10.004

Sapir, E. (1921). *Language. An Introduction to the Study of Speech*. New York: Harcourt, Brace & Co. https://doi.org/10.1016/j.livsci.2008.10.015

Schell, M., Zaccarella, E., & Friederici, A. D. (2017). Differential cortical contribution of syntax and semantics: An fMRI study on two-word phrasal processing. *Cortex*, *96*, 105–120. https://doi.org/10.1016/J.CORTEX.2017.09.002

Seidenberg, M., & MacDonald, M. (1999). A probabilistic constraints approach to language acquisition and processing. *Cognitive Science*, *23*(4), 569–588.

Tai, J. H.-Y. (2005). Conceptual Structure and Conceptualizations in Chinese Grammar. *Language and Linguistics*, *6*(4), 539–574.

Tai, J. H.-Y. (2013). Reflections on Typological Characterization of Chinese Grammar. *Human Language Resources and Linguistic Typology*, 59–88.

Thurgood, G., & LaPolla, R. J. (2003). *The Sino-Tibetan languages*. Routledge.

Townsend, D. J., Bever, T. G., & Crocker, M. W. (2001). *Sentence Comprehension : The Integration of Habits and Rules*. *Computational Linguistics* (Vol. 28).

van der Burght, C. L., Goucha, T., Friederici, A. D., Kreitewolf, J., & Hartwigsen, G. (2019). Intonation guides sentence processing in the left inferior frontal gyrus. *Cortex*, *117*, 122–134. https://doi.org/10.1016/j.cortex.2019.02.011

Wu, C.-Y., Zaccarella, E., & Friederici, A. D. (2018). Universal neural basis of structure building evidenced by network modulations emerging from Broca's area: The case of Chinese. *Human Brain Mapping*. https://doi.org/10.1002/hbm.24482

Xu, J., Kemeny, S., Park, G., Frattali, C., & Braun, A. (2005). Language in context:

Emergent features of word, sentence, and narrative comprehension. *NeuroImage*, *25*(3), 1002–1015. https://doi.org/10.1016/j.neuroimage.2004.12.013

Yarkoni, T., Speer, N. K., & Zacks, J. M. (2008). Neural substrates of narrative comprehension and memory. *NeuroImage*, *41*(4), 1408–1425. https://doi.org/10.1016/J.NEUROIMAGE.2008.03.062

Ye, Z., Luo, Y., Friederici, A. D., & Zhou, X. (2006). Semantic and syntactic processing in Chinese sentence comprehension: Evidence from event-related potentials. *Brain Research*, *1071*(1), 186–196. https://doi.org/10.1016/j.brainres.2005.11.085

Yu, J., & Zhang, Y. (2008). When Chinese semantics meets failed syntax. *Neuroreport*, *19*(7), 745–749. https://doi.org/10.1097/WNR.0b013e3282fda21d

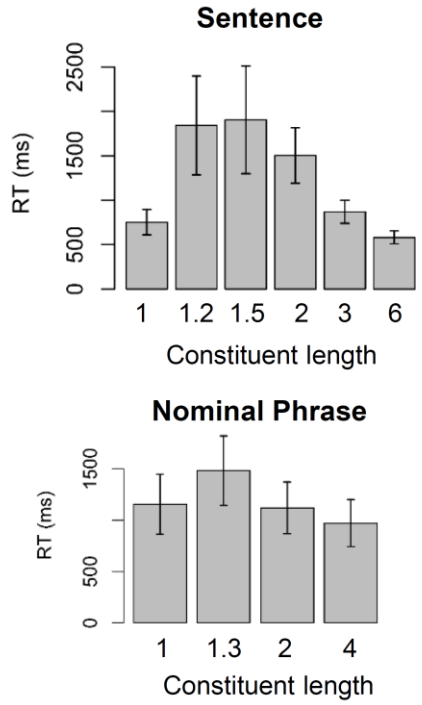
Zaccarella, E., Schell, M., & Friederici, A. D. (2017). Reviewing the functional basis of the syntactic Merge mechanism for language: A coordinate-based activation likelihood estimation meta-analysis. *Neuroscience & Biobehavioral Reviews*, *80*, 646–656. https://doi.org/10.1016/J.NEUBIOREV.2017.06.011

Zhang, Y., Li, P., Piao, Q., Liu, Y., Huang, Y., & Shu, H. (2013). Syntax does not necessarily precede semantics in sentence processing: ERP evidence from Chinese. *Brain and Language*, *126*(1), 8–19. https://doi.org/10.1016/j.bandl.2013.04.001

Zhang, Y., Yu, J., & Boland, J. E. (2010). Semantics does not need a processing license from syntax in reading Chinese. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *36*(3), 765–781. https://doi.org/10.1037/a0019254

Zhou, X., Ye, Z., Cheung, H., & Chen, H.-C. (2009). Processing the Chinese language: An introduction. *Language and Cognitive Processes*, *24*(7–8), 929–946. https://doi.org/10.1080/01690960903201281

Appendix. Results of constituent assessment. Right: number of stimuli reported in all the possible constituent lengths. Left: Assessment time for stimuli of different reported constituent lengths. Error bars indicate 95% CI. Constituent lengths not included in our original design (length 1.2, 1.5, and 2 for sentence condition and length 1.3 for nominal phrase condition) were sparsely reported and had longer RTs. Therefore, they were modeled with a separate regressor in the individual fMRI model and excluded from the group level model.



Subject ID	Sentence Reported constituent length						Nominal Phrase Reported constituent length			
	1	1.2	1.5	2	3	6	1	1.3	2	4
2	35	0	0	0	24	31				
3	30	0	0	0	27	33				
4	34	2	1	2	16	31				
5	19	0	7	4	29	31				
6	36	0	0	1	24	29				
7	34	1	6	8	18	23	22	10	16	12
10	30	4	2	4	26	24	20	3	21	16
11	31	0	0	0	27	32	17	2	18	23
12	28	3	0	5	23	31	27	7	19	7
13	25	0	7	4	25	29	22	13	19	6
14	29	1	0	1	23	36	11	1	21	27
15	17	3	0	4	31	35	12	2	28	18
16	22	4	0	8	19	37	32	2	7	19
17	28	0	1	3	26	32	26	0	21	13
18	27	2	2	0	28	31	15	19	22	4
19	28	2	2	6	24	28	11	13	29	7
20	30	0	1	1	27	31	17	10	20	13
21	32	0	0	1	29	28	23	0	30	7
22	25	4	3	6	23	29				
23	21	3	1	5	30	30	17	6	22	15
25	24	4	0	6	23	33				
26	29	0	0	3	23	35	17	11	25	7