Brain & Language 229 (2022) 105110

Contents lists available at ScienceDirect

Brain and Language

journal homepage: www.elsevier.com/locate/b&l

Neural correlates of semantic number: A cross-linguistic investigation

Donald Dunagan [a,*], Shulin Zhang [a], Jixing Li [b], Shohini Bhattasali [c], Christophe Pallier [d], John Whitman [e], Yiming Yang [f], John Hale [a]

[a] Department of Linguistics, University of Georgia, GA, USA
[b] Neuroscience of Language Lab, New York University Abu Dhabi, Abu Dhabi, United Arab Emirates
[c] Department of Linguistics, UMIACS, University of Maryland, MD, USA
[d] Cognitive Neuroimaging Unit, INSERM-CEA, Gif-sur-Yvette, France
[e] Department of Linguistics, Cornell University, NY, USA
[f] Jiangsu Normal University, Xuzhou, Jiangsu, China

ARTICLE INFO

Keywords:
Neurolinguistics
Semantics
Typology
Number
Cognitive neuroscience

ABSTRACT

One aspect of natural language comprehension is understanding how many of what or whom a speaker is referring to. While previous work has documented the neural correlates of number comprehension and quantity comparison, this study investigates semantic number from a cross-linguistic perspective with the goal of identifying cortical regions involved in distinguishing plural from singular nouns. Three fMRI datasets are used in which Chinese, French, and English native speakers listen to an audiobook of a children's story in their native language. These languages are selected because they differ in their number semantics. Across these languages, several well-known language regions manifest a contrast between plural and singular, including the pars orbitalis, pars triangularis, posterior temporal lobe, and dorsomedial prefrontal cortex. This is consistent with a common brain network supporting comprehension across languages with overt as well as covert number-marking.

1. Introduction

One aspect of natural language comprehension is understanding how many of what or whom a speaker is referring to. While much work has been done to document the neural correlates of representing, manipulating, and comparing numerical quantities (e.g., Castelli, Glaser, & Butterworth, 2006; Dehaene, Piazza, Pinel, & Cohen, 2003; Kadosh & Walsh, 2009), the neural correlates of nominal plurality are less well understood. This paper seeks to help remedy this by investigating brain activity associated with the comprehension of plural and singular nouns while participants listen to a story. French, Mandarin Chinese, and English have differences in their number semantics which result in differences in their morpho-syntactic marking of number. For this reason, the results of separate, but parallel fMRI analyses using these languages are compared to see if typological differences result in differences in neural activity.

The term *semantic number* (as compared to *grammatical number*) is used throughout the paper because, while grammatical number is a grammatical category which is expressed either through morphology or syntax, not all of the languages which are analyzed mark number morpho-syntactically. More detail regarding terminological definitions and the semantics of number is given in Section 1.2.

Cross-linguistically, languages with singular/plural nominal contrast tend to mark plural forms morphologically, but not mark singular forms (Corbett, 2000; Greenberg, 1963). In the semantics literature, Farkas and de Swart (2010) propose a weak singular/strong plural account in which plurality is morphologically and semantically marked and singularity is not. A short review of the psycholinguistics of number in Section 1.1 connects this singular/plural asymmetry to existing behavioral work. Section 1.2 introduces a line of semantic typology work which motivates the choice of languages in this study. Related topics such as arithmetic number processing and the domain-general versus language-specific distinction are treated in Section 1.3.

1.1. *The psycholinguistics of number*

Psycholinguistic research has shown a behavioral asymmetry between nominal singularity and plurality during production and

* Corresponding author at: Department of Linguistics, 142 Gilbert Hall, University of Georgia, Athens, GA 30602-6205, USA.
E-mail address: dgd45125@uga.edu (D. Dunagan).

https://doi.org/10.1016/j.bandl.2022.105110
Received 12 May 2021; Received in revised form 22 March 2022; Accepted 24 March 2022
Available online 31 March 2022

comprehension. There exists a plural complexity effect such that plural nouns are more difficult to process than singular nouns. On the basis of number agreement error patterns in production, Eberhard (1997) proposes a difference in feature marking. During comprehension, the asymmetry is measurable, for example, in differences in P600s (Tanner, Nicol, & Brehm, 2014). We expect this plural complexity effect to be observable with fMRI. That is, because they are harder to process, we expect that plural nouns will elicit greater cortical activity than singular nouns.

In production, the asymmetry has emerged in experiments that elicit agreement errors. As an example, in the erroneous sentence, "The cost of the improvements have not yet been estimated" (Bock & Miller, 1991, p. 1), the auxiliary agrees in number with the immediately preceding, local noun phrase ("the improvements"; NP) rather than the more distant head noun ("cost"). Quirk, Greenbaum, Leech, and Svartvik (1972) give the term "proximity concord" for these errors while Zandvoort (1961) calls them "attraction errors." When participants complete English sentence fragments which contain head NPs and local NPs which disagree in number, Bock and Miller (1991) find that rate of error is greater when the local NP is plural than when the local NP is singular. This number agreement attraction asymmetry has been replicated in other English studies (e.g., Bock & Cutting, 1992; Bock & Eberhard, 1993; Eberhard, 1997) as well as in studies in other languages which mark number, for example, French (Fayol, Largy, & Lemaire, 1994) and German and Dutch (Hartsuiker, Schriefers, Bock, & Kikstra, 2003).

The explanation proposed by Eberhard (1997) for the asymmetry is that singular count nouns are unmarked with respect to number (i.e., they do not have a grammatical feature for number), plural count nouns are marked for number (i.e., they possess a feature for number), and the singular verbal form, rather than being triggered by the presence of a singular noun, is a default which can be overridden by the presence of a plural noun. While a local plural noun may disrupt the number agreement process because it has an activated number feature which could trigger the retrieval of a plural verbal form, a local singular noun is less likely to disrupt the number agreement process because it does not have an activated number feature to be detected. Eberhard, Cutting, and Bock (2005) extend this notion with a computational model focused around SAP (Singular-And-Plural), a real-valued feature for marking number. Negative values and zero indicate singularity and positive values indicate plurality. Total SAP for an NP is calculated, in part, based upon the SAP values of its components according to a weighted, spreading activation-like processes. The attraction error asymmetry follows from this algebraic account.

There is also evidence from word-by-word comprehension that plurals are harder. For grammatical sentences, Pearlmutter, Garnsey, and Bock (1999) find slower reading times in sentences with singular head NPs and plural local NPs than for sentences with singular head NPs and singular local NPs. They interpret this as a sensitivity to number (mis) match between head and local NPs and to an increase in processing difficulty. With their reading time experiments, Wagers, Lau, and Phillips (2009) provide evidence against grammatical sentences being subject to attractor effects, and instead attribute increases in reading times following plural attractors to an increased processing cost for plural nouns. They argue that the attraction effects which they observe in ungrammatical sentences—plural attractors causing an erroneously plural verb to be considered grammatical—result from retrieval errors of a cue-based retrieval system. Extending this line of investigation to event-related brain potentials, Tanner et al. (2014) find that, in ungrammatical sentences, plural attractors produce smaller P600s than singular attractors, but in grammatical sentences, plural attractors have no effect. Additionally, they find slower response times and less accurate (but not quite significantly so) judgments for grammatical sentences with plural attractor nouns in a timed sentence judgment task. They attribute this not to agreement processing, but an increased processing cost for complex NPs which contain plural attractors. They conclude that the causes of number agreement attraction in comprehension are not the same as those in production: while agreement attraction in production seems to involve the number feature of the subject NP, agreement attraction in comprehension seems to involve interference by morphosyntax during the process of cue-based memory retrieval.

With a self-paced reading experiment, Tucker, Idrissi, and Almeida (2015) find that Modern Standard Arabic stimuli can be used to elicit attraction error results similar to those of Wagers et al. (2009). They also find that plural NPs created through suffixation have longer reading times than singular NPs while plural NPs created through ablaut do not. The authors interpret this as a "plural complexity effect" (plural NPs are more difficult to process than singular NPs) as described by Wagers et al. (2009) rather than a "plural integration effect" (plural NPs are more difficult to integrate into a context with other singular NPs) as described by Tanner et al. (2014).

The semantics literature and the psycholinguistics of number literature are in agreement in proposing an asymmetrical complexity effect between singular and plural nouns. However, it remains unknown whether this effect will extend to fMRI. The first question this study addresses is: "As measured with fMRI, do plural nouns elicit greater cortical activity than singular nouns?".

1.2. The choice of these languages

Number in language is made a more interesting topic because languages can differ in their number semantics. Studying Mandarin Chinese, French, and English allows this study to span the full range of values in an important semantic typology. This typology suffices to derive the number marking and non-marking patterns in these languages. The following section reviews the semantics which underlie this three-way contrast.

This paper operates under a simplistic account of the semantics of number in which singular nominals refer to individuals and plural nominals refer to sets of individuals (but see Farkas & de Swart, 2010, Rullmann, 2002, for more detailed accounts). In the proposal by Link (1983), the entity domain to which nominals refer is a join-semilattice. Atoms are individuals and the non-atomic elements are the possible sums of multiple atoms. In these terms, singular nominals choose referents from the domain of atoms and plural nominals chose referents from the domain of sums. Nominals with general number are also considered (Corbett, 2000). These are nominals which are "neutral" or "unspecified" for number. Rullmann and You (2006) describe a system for languages with general number, like Mandarin Chinese, in which atoms generate a complete semilattice and nouns choose referents from the domain of atoms and sums. It is important to note that nominals with general number are not ambiguous between singular and plural readings (see Rullmann & You, 2006, for additional discussion). Their number interpretation might best be given in English as, "one or more *X.*".

A count noun (e.g., *cat*) is a noun which may be directly modified by a cardinal numerical and a mass noun (e.g., *sand*) is a noun which cannot. While (1a) is perfectly acceptable, (1b) is not acceptable on the intended reading. There is a connection between the count/mass distinction and the counting/measuring distinction. While count nouns are counted (1a), mass nouns are measured (1c). It is not the case, however, that all languages make the count/mass distinction.

(1)	a.	two cats
	b.	#two sands
	c.	three buckets of sand

Chierchia (1998) proposes the Nominal Mapping Parameter, which creates a three-way typological classification for languages based upon how they express counting. Chierchia's account is neo-Carlsonian, that is, it is based upon Carlson (1977)'s (Carlson (1977)) investigation of bare plurals in English. Bare nouns are nouns which occur without a determiner or a classifier. This account proposes that nouns can either be predicates at type $<e,t>$, in which they denote a set of entities, or be arguments at type *e*, in which they denote kinds. The terms *predicate* and *argument*, here, are names for the semantic types <*e*,*t*> (functions from individuals to truth values) and *e* (entities of argumental type), respectively. Kinds are generally understood as regularities. For the property of being a cat, there is a corresponding kind: the cat-kind. In the other direction, a kind will have a property with which it corresponds: the property of belonging to the kind.

A noun (N) may fill an argument position if it is an argument, but if it is a predicate, it must combine with a determiner to reach the argument type. Chierchia's classification, then, is whether nouns in a language can occur as arguments, predicates, or both. From the features [+/-predicate] and [+/-argument], there are three possible language types: [+predicate, +argument], [-predicate, +argument], and [+predicate, -argument]. The type [-predicate, -argument] is not valid. English is [+predicate, +argument], Chinese is [-predicate, +argument], and French is [+predicate, -argument]. Chierchia argues that a language will have morphosyntactic properties based upon its features. The following section reviews these properties with data from Rothstein (2017, pp. 147–148, here, Examples 2–5).

With English being [+predicate, +argument], the nouns of English are either [+predicate] or [+argument]. Count nouns are predicates and mass nouns are arguments. Because they are predicates, singular count nouns must combine with a determiner to fill an argument position and it is predicted that bare singular count nouns are ungrammatical (2a). Plural count nouns can be shifted such that they yield a kind reading and thus can occur as bare arguments. Mass nouns can occur bare in argument position (2b).

(2)	a.	I saw #(a) dog.
	b.	I bought wine.

Chinese allows for NPs consisting of bare nouns without classifiers, number morphemes, or other functional elements. Like other classifier languages, Chinese is [-predicate, +argument]. In these languages, bare nouns can occur as arguments (3). While nouns may occur bare, they may not be directly modified by cardinal numericals. Instead of directly taking bare nouns as complements, numericals take classifier (Cl) + N sequences (4).

(3)	a.	wǒ	kànjiàn	gǒu le.
		I	see	dog PARTICLE
		'I saw a dog/dogs, the dog(s).'		
	b.	wǒ	mǎi le	jiǔ.
		I	buy PERFECTIVE	wine
		'I bought wine.'		

(4)	a.	sān	#(zhī)	gǒu
		three	$Cl_{small\ animal}$	dog
		'three dogs'		
	b.	liǎng	#(kē)	shù
		two	Cl_{plant}	tree
		'two trees'		

In an analysis of bare noun phrases in Chinese, Yang (2001) identifies the same narrowest-scope indefinite, kind, and generic readings that Carlson (1977) identifies for English bare plurals (Dayal & Sağ, 2020). While French and English necessarily mark definite NPs with determiners, Chinese does not have determiners and bare NPs have definite readings that are not available in English (Dayal & Sağ, 2020). Since all nouns have the same properties, and no N can be directly modified by a numeral, there is no clear way to differentiate mass and count nouns grammatically (Rothstein, 2017). As compared to languages with mass/count distinction ([+predicate, +/- argument]), Chierchia (1998)'s (Chierchia (1998)) view is that in [-predicate, +argument] languages, every lexical noun is mass-like. Because the plural operator does not apply to kind or mass terms, classifier languages do not have nominal pluralization (Rothstein, 2017). Bare nouns in these languages have a number interpretation which is general and includes the plural (3a).

French, like other Romance languages, is [+predicate, -argument] and makes the count/mass distinction. Count nouns will be marked either singular or plural and all nouns (both count and mass) must occur with a determiner (5).

(5)	a.	J'ai	vu #(un)	chien.
		I AUXILIARY	saw a	dog
		'I saw a dog.'		
	b.	J'ai	acheté #(du)	vin.
		I AUXILIARY	bought some	wine
		'I bought (some) wine.'		

Interestingly, French is slightly more strict with its determiner requirement than Spanish and Italian which allow for bare plurals in well-governed conditions such as object (but not subject) position. The allowed bare plurals do not have generic or kind readings, though (Dayal, 2011). Because English is [+predicate, +argument], its count nouns are similar to French nouns and its mass nouns are similar to Chinese nouns.

1.3. Additional considerations

1.3.1. Number sense

It is possible that semantic number processing is subserved by the same system which subserves human *number sense*, "a short-hand for our ability to quickly understand, approximate, and manipulate numerical quantities" (Dehaene, 2001, p. 16). Natural numbers are thought to be represented as analog magnitudes along a mental number line. Dehaene (1992) proposes a tripartite account of number sense in the human brain. In this *triple-code* model, three portions of the parietal lobe perform different roles in number processing (Dehaene et al., 2003). The horizontal segment of the intraparietal sulcus serves as the location of the mental number line and is augmented by an angular gyrus verbal system and a posterior, superior parietal visual and attentive system.

1.3.2. Language-specific versus domain-general processing

Carreiras, Carr, Barber, and Hernandez (2010) ask if numerical processing is activated by grammatical number processing and, for stimuli with grammatical number violations, find an increase in activation in parietal regions previously implicated in number processing (Dehaene et al., 2003). Portions of the prefrontal and parietal cortices, known as the multiple-demand (MD) network (Duncan, 2010), have been found to be responsive to a wide variety of cognitive demands such as: verbal and spatial working memory, the Stroop task, and potentially relevant here, an arithmetic task (Fedorenko, Duncan, & Kanwisher, 2013). On one hand, these regions have been shown to not track linguistic input as closely as language-selective regions (Blank & Fedorenko, 2017), and Fedorenko, Behr, and Kanwisher (2011) find little or no overlap between cortical regions engaged in high-level linguistic processing and MD regions which respond to general working memory, cognitive control, and potentially relevant here, mental arithmetic. On the other hand, Carreiras et al. (2010) identify a link between number in language and general number in the brain. The next question this study addresses is: "If plural nouns elicit greater cortical activity than singular nouns, do these regions of increased activation align with regions known for quantity and arithmetic processing or with regions that are known for linguistic processing?".

1.3.3. Neural, cross-linguistic similarities and differences

While neural, cross-linguistic differences have been found in domains such as phonological access in a reading task (Paulesu et al., 2000), pitch contour processing (Gandour et al., 2003), and nominal and verbal representation (Li, Jin, & Tan, 2004), similarities have been found for syntactic processing (see Obleser, Meyer, & Friederici, 2011; Pallier, Devauchelle, & Dehaene, 2011, for German & French results, respectively) and comprehending linguistic content (Honey, Thompson, Lerner, & Hasson, 2012). The final question to be addressed is: "Although they differ in their number semantics, if French, English, and Chinese display increased activation for plural nouns over singular nouns, does that activation occur in the same or different regions?".

2. Data and methods

The analyses use *The Little Prince Datasets* (Li et al., 2021), a group of three fMRI datasets collected as French, Chinese, and English speaking participants listen to *Le Petit Prince* (*The Little Prince,* de Saint-Exupéry, 1946), a children's storybook, in their native language. Reflections on this and other neurolinguistic datasets can be found in Hale et al. (2022).

To localize the brain regions involved in the processing of singular and plural nouns during naturalistic language comprehension, separate, whole-brain general linear model (GLM) analyses are performed for each of the three separate languages, computing a PLURAL > SINGULAR contrast. Then, in order to investigate the similarities and differences between the results for the separate languages, the overlap of voxels which were statistically significant in the separate GLM analyses is examined. This is possible because all three datasets are in Montreal Neurological Institute (MNI) space.

2.1. Participants

English participants were 51 healthy, right-handed young adults (32 females, mean age = 21.3, SD = 3.6). They self-identified as native English speakers and had no history of psychiatric, neurological or other medical illness that might compromise cognitive functions. All participants were paid, and gave written informed consent prior to participation, in accordance with the IRB guidelines of Cornell University.

Chinese participants were 35 healthy, right-handed young adults (15 females, mean age = 19.3, SD = 1.6). They self-identified as native Chinese speakers and had no history of psychiatric, neurological, or other medical illness that could compromise cognitive functions. All participants were paid, and gave written informed consent prior to participation, in accordance with the IRB guidelines of Jiangsu Normal University.

French participants were 30 healthy, right-handed young adults (16 females, mean age = 24.3, SD = 4.9). They self-identified as native French speakers and had no history of psychiatric, neurological, or other medical illness that could compromise cognitive functions. All participants gave written informed consent prior to participation, in accordance with the Regional Committee for the Protection of Persons involved in Biomedical Research.

2.2. Experimental procedure and stimuli

After giving their informed consent, participants were familiarized with the MRI facility and assumed a supine position on the scanner. Auditory stimuli were delivered through MRI-safe, high-fidelity headphones inside the head coil. The headphones were secured against the plastic frame of the coil using foam blocks. An experimenter increased the sound volume stepwise until the participants could hear clearly. The French audio stimulus was an audiobook version of *Le Petit Prince* (de Saint-Exupéry, 1946), read by Nadine Eckert-Boulet. The English audio stimulus was an English translation of *The Little Prince*, read by Karen Savage. The Chinese audio stimulus was a Chinese translation of *The Little Prince*, read by a professional female Chinese broadcaster. The stimuli were divided into nine sections, with each lasting for about 10 min. Participants listened passively to the nine sections and completed four quiz questions after each section (36 questions in total). These questions were used to confirm their comprehension and were viewed by the participants via a mirror attached to the head coil. Participants answered through a button box. The entire session, including preparation time and practice, lasted for around 2.5 h.

2.3. Data acquisition and preprocessing

English and Chinese MRI images were acquired with a 3T MRI GE Discovery MR750 scanner with a 32-channel head coil. French MRI images were acquired with a 3T Siemens Magnetom Prisma Fit 230 scanner. Anatomical scans were acquired using a T1-weighted volumetric magnetization prepared rapid gradient-echo pulse sequence. Functional scans were acquired using a multi-echo planar imaging sequence with online reconstruction (TR = 2000 ms; English and Chinese: TEs = 12.8, 27.5, 43 ms; French: TEs = 10, 25, 38 ms; FA = 77°; matrix size = 72 x 72; FOV = 240.0 mm x 240.0 mm; 2 x image acceleration; English and Chinese: 33 axial slices; French: 34 axial slices; voxel size = 3.75 x 3.75 x 3.8 mm). All fMRI data were preprocessed using AFNI version 16 (Cox, 1996). The first 4 volumes in each run were excluded from analyses to allow for T1-equilibration effects. Multi-echo independent components analysis (Kundu, Inati, Evans, Luh, & Bandettini, 2012) was used to denoise data for motion, physiology and scanner artifacts. Images were then spatially normalized to the standard space of the MNI atlas, yielding a volumetric time series resampled at 2 mm cubic voxels for the English and Chinese data and 3.15 mm cubic voxels for the French data.

2.4. Observations of interest

In order to control for discourse factors which could modulate neural activity during naturalistic language processing, the storybook texts are aligned and only parallel nouns are selected for analysis. That is, nouns which occur in all three stories and in the same context. The first step in this process is aligning sentences, which is done with the Hunalign bilingual sentence aligner (Varga et al., 2005). The alignments were checked and corrected by hand. Next, the parallel nouns are identified and filtered with criteria which serve to maximize the typological contrast between French, English, and Chinese nominals.

For the Chinese observations, we include only nouns which have no overt number marking, either morphological or through a number and classifier construction. This captures the [+argument] aspect of Chinese. For the French observations, we include only definite count nouns indexed by the definite, common determiners: *le*, *la*, *l'*, and *les*. This captures the [+predicate] aspect of French and its requirement that definite nouns be marked with a definite determiner. For the English observations, we include only count nouns, but allow them to be definite, indefinite, or type-shifted bare plurals. This capture the [+predicate, +argument] aspects of English.

While number annotation can be automated for the French nouns: *le*, *la*, and *l'* are singular and *les* is plural, and English count nouns are easily annotated for number based upon their overt number morphology, annotation for the Chinese nouns is more challenging because number is not overtly marked. Recall that Chinese bare nominals are not ambiguous between singular and plural readings, but it is possible that different listeners will have different judgements. Because of this, we have two native Chinese speakers annotate the Chinese nouns with singular/plural judgments. Calculating Cohen's kappa coefficient (Cohen, 1960), a measure of inter-rater reliability, over the Chinese annotations results in a kappa=0.96, a high degree of inter-rater reliability. Nouns for which the two annotators disagreed in their number judgements were not included in the analysis.

The time resolution for all three of the fMRI data sets is 2.0 seconds, much slower than a natural speech rate. Because of this, observations where nouns of different number would occur together within the same volume were removed. That is, if more than one singular noun occur in the same volume or if more than one plural noun occur in the same volume, they are retained. If a singular and plural noun occur in the same volume, however, they are not kept for analysis. This selection results in 274 parallel observations: 245 singular and 29 plural in the Chinese text, 245 singular and 29 plural in the French text, and 244 singular and 30 plural in the English text.

2.5. Statistical analyses

All analyses are carried out using Nilearn (Abraham et al., 2014; Pedregosa et al., 2011, version 0.7.1), a package for the statistical analysis of neuroimaging data in Python. At the first level of the GLM, voxel blood-oxygen-level-dependent (BOLD) timecourses are modeled: the observed BOLD time series is the dependent variable and the independent variables are the time series of the stimuli convolved with the hemodynamic response function. Binary regressors are included for the singular and plural noun observations of interest from the storybooks, timelocked to their offsets. Coregressors of non-interest are also included for spoken word rate, log lexical frequency, root mean squared amplitude of the spoken narration (RMS), and speaker pitch. Speaker pitch is not available for the French audiobook, but the other coregressors are used. The coregressors are added to ensure that any results found are due to the differences between singular and plural nouns and not just effects of spoken language comprehension (cf. Bullmore et al., 1999; Lund, Madsen, Sidaros, Luo, & Nichols, 2006).

The singular and plural noun regressors are marked with a 1 at the offset of the nouns of interest, word rate is marked with a 1 at the offset of every word, except for the observations of interest, log lexical frequency is marked at the end of every word, and RMS and pitch are marked every 10 ms.

Following the unmarked singular/marked plural semantic account of number proposed by Farkas and de Swart (2010) (at least for French and English which make this distinction), as well as the "plural complexity effect," where plural NPs are more difficult to process than singular NPs, from the psycholinguistics of comprehension (Tucker et al., 2015; Wagers et al., 2009), the first level contrasts use the fitted first level GLMs to subtract activity associated with singular nouns away from activity associated with plural nouns. That is, a PLURAL > SINGULAR contrast is computed.

At the second level of the GLM analyses, the first level contrast maps are used to perform one-sample t-tests: "Is the difference between plural and singular activation at this voxel greater than 0?" An 8 mm full width at half maximum Gaussian smoothing kernel is applied to counteract inter-subject anatomical variation. The by-language, group-level results reported in the following section underwent family-wise-error (FWE) voxel correction for multiple comparisons and are reported in terms of z-score. Only clusters greater than 100 mm^3 are retained. Additionally, the overlap of the separate results are reported in order to identify any common regions of increased activation between the three languages. The MNI2TAL tool from the Yale BioImage Suite (Lacadie, Fulbright, Arora, Constable, & Papademetris, 2008a; Lacadie, Fulbright, Rajeevan,

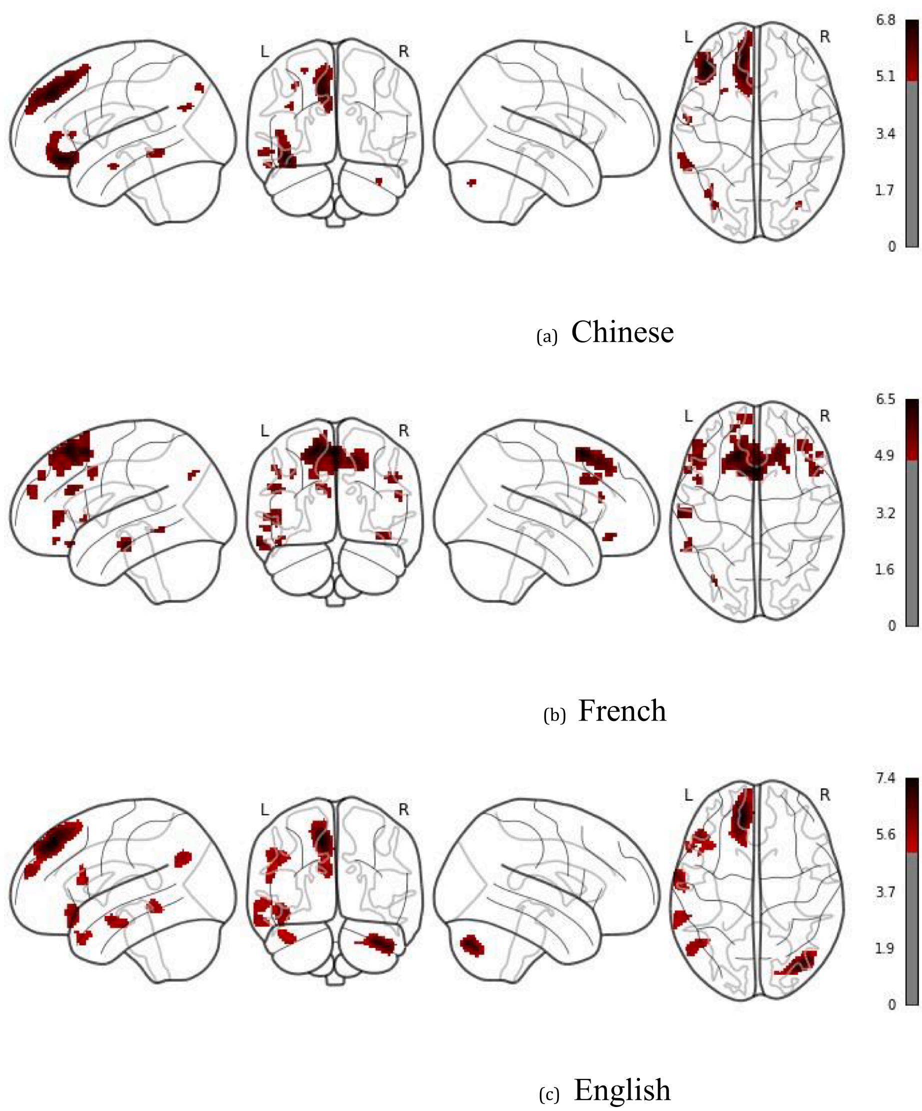

Fig. 1. Significant clusters for the PLURAL > SINGULAR contrast (z-valued) after FWE voxel correction for multiple comparisons with $p < 0.05$, cluster size > 100 mm^3 for Chinese (a), French (b), and English (c).

Constable, & Papademetris, 2008b,version 1.4) was referenced for brain region and Brodmann area labels.

3. Results

3.1. Chinese results

For the Chinese participants, an increase in activation is found for plural nouns over singular nouns in the left pars triangularis (BA 45) and left pars orbitalis (BA 47), extending into left dorsolateral prefrontal corex (BA 46), in left dorsomedial prefrontal cortex (BAs 8, 9), the left fusiform (BA 37), angular (BA 39), and middle temporal (BA 21) gyri, and the right cerebellum. These results can be seen in Fig. 1a and more detail can be found in Table 1.

3.2. French results

For the French participants, an increase in activation is found for plural nouns over singular nouns in the left and right pars opercularis (BA 44), left pars triangularis (BA 45), left and right pars orbitalis (BA 47), left (BAs 8, 9, 10) and right (BA 8) dorsomedial prefrontal cortex, left (BAs 8, 46) and right (BA 9) dorsolateral prefrontal cortex, and the left middle temporal (BA 21), fusiform (BA 37), and angular (BA 39) gyri. These results can be seen in Fig. 1b and more detail can be found in Table 2.

3.3. English results

For the English participants, an increase in activation is found for plural nouns over singular nouns in the left pars opercularis (BA 44), left pars triangularis (BA 45), left pars orbitalis (BA 47), left dorsomedial prefrontal cortex (BAs 8, 10), the left temporal pole (BA 38), the left middle temporal (BA 21) and angular gyri (BA 39), and the right cerebellum. These results can be seen in Fig. 1c and more detail can be found in Table 3.

3.4. Cross-linguistic overlap

Overlaying the significant clusters from the Chinese, French, and English main results, voxel-wise overlap between all three languages is found in the left pars orbitalis (BA 47) and left dorsomedial prefrontal cortex (BAs 8, 10), as indicated in black in Fig. 2. Voxel-wise overlap between two languages is found in the left pars orbitalis (BA 47), left dorsomedial prefrontal cortex (BA 8), left dorsolateral prefrontal cortex (BA 46), the left middle temporal (BA 21) and fusiform (BA 37) gyri, and the right cerebellum, as indicated in red in Fig. 2. More detail can be found in Table 4. Additionally, while voxel-level overlap is not observed, all three languages show an increase in activation in the left pars triangularis (BA 45), and the left angular gyrus (BA 39).

Table 1
Significant PLURAL > SINGULAR clusters for Chinese after FWE voxel correction for multiple comparisons with $p < 0.05$, cluster size > 100 mm^3.

Region	Cluster size (mm^3)	MNI coordinates x	y	z	Peak stat (z)
L Pars Orbitalis (BA 47)	4936	−40.0	38.0	−16.0	6.85
L Dorsolateral Prefrontal Cortex (BA 46)		−46.0	42.0	2.0	5.62
L Dorsomedial Prefrontal Cortex (BA 9)	6952	−12.0	44.0	38.0	6.70
L Dorsomedial Prefrontal Cortex (BA 8)		−10.0	30.0	48.0	6.45
		−8.0	20.0	56.0	5.64
L Fusiform Gyrus (BA 37)	768	−60.0	−42.0	−12.0	5.54
L Angular Gyrus (BA 39)	304	−36.0	−64.0	26.0	5.54
L Middle Temporal Gyrus (BA 21)	192	−56.0	−6.0	−22.0	5.44
L Pars Triangularis (BA 45)	208	−46.0	28.0	4.0	5.39
R Cerebellum	136	36.0	−76.0	−36.0	5.39
L Angular Gyrus (BA 39)	176	−34.0	−78.0	44.0	5.33
L Dorsal Prefrontal Cortex (BA 8)	104	−26.0	18.0	56.0	5.21

Table 2
Significant PLURAL > SINGULAR clusters for French after FWE voxel correction for multiple comparisons with $p < 0.05$, cluster size > 100 mm^3.

Region	Cluster size (mm^3)	MNI coordinates x	y	z	Peak stat (z)
Dorsomedial Prefrontal Cortex (BA 8)	14917	−14.0	30.0	60.0	6.48
		2.0	15.0	56.0	6.45
		−8.0	18.0	56.0	6.34
		17.0	30.0	50.0	5.91
L Middle Temporal Gyrus (BA 21)	1040	−65.0	−17.0	−19.0	6.07
R Pars Opercularis (BA 44)	914	52.0	18.0	34.0	5.66
R Dorsolateral Prefrontal Cortex (BA 9)		46.0	27.0	34.0	5.25
L Pars Opercularis (BA 44)	1040	−49.0	27.0	25.0	5.62
L Dorsolateral Prefrontal Cortex (BA 46)	946	−49.0	40.0	3.0	5.43
R Pars Orbitalis (BA 47)	630	39.0	40.0	−13.0	5.38
L Dorsolateral Prefrontal Cortex (BA 8)	473	−49.0	11.0	41.0	5.29
R Dorsolateral Prefrontal Cortex (BA 9)	220	55.0	30.0	18.0	5.27
L Dorsomedial Prefrontal Cortex (BA 10)	693	−8.0	62.0	25.0	5.26
L Pars Orbitalis (BA 47)	157	−46.0	30.0	−19.0	5.23
L Dorsomedial Prefrontal Cortex (BA 9)	378	−17.0	56.0	37.0	5.21
L Fusiform Gyrus (BA 37)	378	−58.0	−45.0	−10.0	5.13
L Pars Triangularis (BA 45)	189	−55.0	18.0	−0.0	5.04
L Angular Gyrus (BA 39)	220	−33.0	−74.0	41.0	5.03
L Pars Orbitalis (BA 47)	157	−49.0	43.0	−19.0	5.01

Table 3
Significant PLURAL > SINGULAR clusters for English after FWE voxel correction for multiple comparisons with $p < 0.05$, cluster size > 100 mm^3.

Region	Cluster size (mm^3)	MNI coordinates x	y	z	Peak stat (z)
L Dorsomedial Prefrontal Cortex (BA 8)	9176	−12.0	44.0	46.0	7.43
L Dorsomedial Prefrontal Cortex (BA 10)		−8.0	64.0	22.0	6.00
R Cerebellum	3232	36.0	−80.0	−38.0	7.04
		44.0	−74.0	−42.0	6.89
		18.0	−86.0	−38.0	5.18
L Pars Orbitalis (BA 47)	1616	−44.0	28.0	−14.0	6.22
L Angular Gyrus (BA 39)	1512	−52.0	−64.0	34.0	5.96
L Middle Temporal Gyrus (BA 21)	1272	−66.0	−6.0	−18.0	5.94
		−62.0	−14.0	−22.0	5.63
	1088	−62.0	−38.0	−8.0	5.76
L Temporal Pole (BA 38)	728	−38.0	20.0	−36.0	5.52
		−46.0	16.0	−32.0	5.31
L Pars Opercularis (BA 44)	648	−54.0	20.0	22.0	5.48
L Pars Triangularis (BA 45)		−56.0	18.0	14.0	5.47

4. Discussion

In contrasting neural activation between plural and singular nouns, several common regions of increased activation are found: the left pars orbitalis (POrb), left pars triangularis (PTri), left dorsomedial (DMPFC) and dorsolateral (DLPFC) prefrontal cortex, the left middle temporal (MTG), fusiform, and angular gyri (AG), and the right cerebellum.

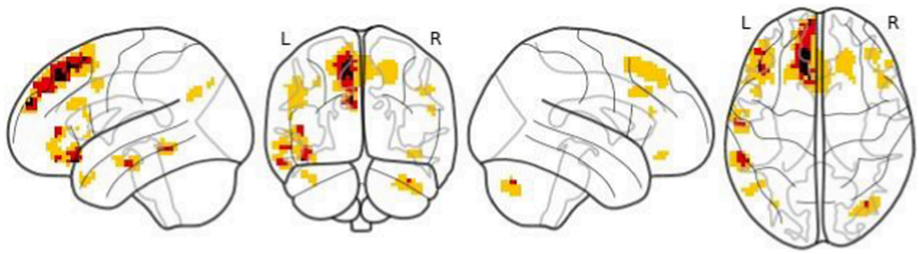

Fig. 2. Overlap of Chinese, French, and English main results. Yellow indicates significant PLURAL > SINGULAR clusters for 1 language, red for 2 languages, and black for all 3 languages.

Table 4
Clusters resulting from overlap of Chinese, French, and English PLURAL > SINGULAR main results. Only clusters where 2 or more languages overlap are presented.

Region	Cluster size (mm^3)	MNI coordinates			Overlapping languages
		x	y	z	
L Pars Orbitalis (BA 47)	504	−46.0	27.0	−19.0	3.0
L Dorsomedial Prefrontal Cortex (BA 10)	504	−8.0	62.0	25.0	3.0
L Dorsomedial Prefrontal Cortex (BA 8)	4194	−14.0	30.0	53.0	3.0
		−11.0	40.0	47.0	3.0
		−8.0	24.0	56.0	3.0
		−21.0	30.0	53.0	2.0
L Middle Temporal Gyrus (BA 21)	189	−62.0	−14.0	−23.0	2.0
L Fusiform Gyrus (BA 37)	220	−58.0	−42.0	−10.0	2.0
L Dorsolateral Prefrontal Cortex (BA 46)	220	−46.0	43.0	3.0	2.0
L Pars Orbitalis (BA 47)	126	−46.0	40.0	−16.0	2.0
L Dorsomedial Prefrontal Cortex (BA 8)	31	−5.0	27.0	50.0	2.0
R Cerebellum	63	36.0	−74.0	−35.0	2.0

Voxel-level overlap is observed between all three languages in the left POrb and left DMPFC. These findings do not align with what would be expected if semantic number were subserved by the same brain network as triple-code number processing and quantity comparison (Dehaene et al., 2003). Instead, subsets of these regions have previously been implicated in the syntactic (Flick & Pylkkänen, 2020; Matchin & Hickok, 2020; Murphy, 2020) and semantic (Binder, Desai, Graves, & Conant, 2009; Graessner, Zaccarella, & Hartwigsen, 2021; Graves, Binder, Desai, Conant, & Seidenberg, 2010) processing of language.

4.1. Quantity comparison

If semantic number were just another facet of the human number sense, then we would expect to identify the horizontal segment of the intraparietal sulcus (HIPS). The triple-code model (Dehaene, 1992) proposes a tripartite account for making sense of numbers and quantities with three portions of the parietal lobe facilitating this in different manners (Dehaene et al., 2003). The HIPS serves as an internal number line which keeps track of size and distance between numbers and which is responsible for number representation, the left AG aids in processing heard numbers without processing quantities directly, and the posterior, superior parietal lobe (PSPL) orients attention both in space and on the internal number line. While the HIPS would be a plausible candidate for the PLURAL > SINGULAR contrast, a significant difference in activation is not observed there. Even though Carreiras et al. (2010) observe an increase in activation for grammatical number disagreement in the right HIPS and right PSPL, they believe it unlikely that the parietal regions are specifically involved when processing language and it more likely that the activation is from quantity computation engaged by the grammatical judgement task. Indeed, they perform SINGULAR > PLURAL and PLURAL > SINGULAR contrasts for the identified parietal regions with their grammatical stimuli, but find no effect. While the left AG is identified in the contrast of plural and singular nouns, importantly, the HIPS is not. In this regard, we conclude that the comprehension of semantic number is localized to the language network.

4.2. Semantic processing

One interpretation of these results is that we are observing an effect for semantic number in semantic comprehension. That is, it is either more difficult to integrate plural nouns into the current, working semantic representation than singular nouns, or it is more difficult to mentally represent semantic scenes with plural entities than scenes with singular entities. It is reasonable that whether there are *one* or *many* of someone or something would play a role in constructing meaning during language comprehension and that being morphologically (Corbett, 2000; Greenberg, 1963), semantically (Farkas & de Swart, 2010), and psycholinguistically (Tucker et al., 2015; Wagers et al., 2009) marked, plural nominals would elicit greater activation than singular nominals.

For explicit, two-word semantic composition, Graessner et al. (2021) find an increase in activation in the left inferior frontal gyrus (IFG), left DMPFC, bilateral AG, left pMTG, left ATL, right fusiform gyrus, and the right cerebellum. The areas which they identify in this contrast are nearly the same as those found here. Other fMRI studies which have investigated the neural correlates of semantic processing have found similar, but varied results. Graves et al. (2010) find an increase in activation in the right AG, bilateral DMPFC, and bilateral posterior cingulate and precuneus for two-word semantic composition. For semantic comprehension of sentence-length stimuli, Pallier et al. (2011) identify the left ATL, left anterior superior temporal sulcus (STS), and the left temporo-parietal junction while Humphries, Binder, Medler, and Liebenthal (2006) identify bilateral STS, MTG, inferior temporal gyrus, and the AG (the frontal lobe was not included in their analysis).

DMPFC was implicated across all three languages, and our interpretation is that it plays a specific role as host to deep semantic representations. In a recent neurocognitive proposal, Murphy (2020) posits that theta oscillations serve to combine representations and encode linguistic material (see the discussion in Section 4.3). Mas-Herrero and Marco-Pallarés (2016) demonstrate that theta plays a role in integrating the outputs of different cognitive processes, the functional paths of which converge in medial PFC. DMPFC—serving as integrator of lexical-semantic and syntactic representations from other portions of the language network—is well suited for, as Murphy (2020, pp. 73–74) states, “rapidly encoding linguistic material given its natural propensity to vanish from working memory.” Working memory discussion of the results is provided in 4.4. Additional support for this role for DMPFC comes from Frankland and Greene (2020), who find that anterior medial PFC (BA 10) represents narrow semantic roles (those which are verb-specific) during event understanding. This is in contrast to the left-mid superior temporal cortex which represents semantic roles which are broad (those which are verb invariant, e.g., Agent, Patient) and more syntactic in nature.

4.3. Syntactic processing

While voxel-level overlap is only observed in two of the three languages, all three languages demonstrate an increase in activation in the posterior temporal lobe (PTL) for plural observations of interest. This is contextualized in the neuroanatomical model of Matchin and Hickok (2020), where the pMTG subserves syntactic processing during comprehension by creating hierarchical lexical-syntactic structure out of the phonological forms received from the posterior superior temporal gyrus. This role for the PTL in syntactic composition is supported by the MEG results of Flick and Pylkkänen (2020) who find evidence for syntactic composition in the PTL, but not in the ATL or IFG.

Murphy (2020) proposes that gamma oscillations encode basic linguistic phi-features, including Number, and that phrasal construction occurs through feature-set construction (at least partially) in the PTL. While this study does not speak to the oscillatory encoding of the Number feature, it does support Murphy's localization to the PTL. It could be the case that, as a default, singularity is not explicitly encoded as a feature to be bundled, while, as the nondefault, plurality requires explicit encoding (cf. Eberhard, 1997).

4.4. Working memory

The PTri and POrb results are explored in the context of syntactic and semantic working memory, respectively. In the model of Matchin and Hickok (2020), the IFG serves an important role during comprehension, either as a form of syntactic working memory or as a predictor of upcoming syntactic structure. While Broca's area is broadly credited with working memory computations such as retrieval and attention during language use, Matchin (2018) goes further and proposes that different anatomical portions perform these functions for different linguistic representations: the PTri selects for syntactic representations, the POrb selects for semantic representations, and the pars opercularis selects for phonological representations.

Taking again an account of number in which singularity is an unencoded default (Eberhard, 1997), the retrieval or maintenance of an explicitly encoded and nondefault plurality feature would impose increased demand on the syntactic working memory space and operations of the PTri. In the introduction, an operational definition of semantic number was offered in which singular nominals refer to individuals and plural nominals refer to sets of individuals. With the POrb providing working memory space and operations for semantic construction, semantically plural nominals would impose a greater demand on the POrb than singular nominals through the retrieval and/or maintenance of sets of individuals as compared to a single individual.

4.5. Cross-linguistic uniformity of brain bases

The similarities seen between the Chinese, English, and French results are remarkable. In bilinguals, previous research has found overlap between the L1 and L2 regions which subserve lexicosemantic comparison (Crinion et al., 2006; Klein et al., 2006). Honey et al. (2012) expand this to narrative level stimuli, analyzing neural activity for monolingual English speakers and bilingual, Russian native, English L2 speakers in two conditions: listening to a Russian story and listening to an English translation of that story. When the participants listen to the story in their native language, they find a number of areas in common which reliably respond to the content of the narrative: the STS, the AG, the supramarginal gyrus, the IFG, the precuneus, the middle frontal gyrus, and orbitofrontal cortex. These results, like ours, show that neural response patterns can be shared across groups despite differences in linguistic form. Importantly, it is found that plurality conveyed through discourse cues (as in Chinese) elicits a similar response to overtly marked plurality (as in French and English).

Ayyash et al. (2021) use functional data from speakers of 45 different languages to demonstrate that the language network is cross-linguistically consistent in its left lateralization, location, functional selectivity, and functional integration. While their work furthers our ability to say that there are not broad differences in the neural implementations of the human language processor for speakers of different languages, there is still more work to be done investigating whether neural similarities or differences emerge when analyses target key axes along which languages differ. That is what we have done here: we identify a dimension along which languages differ (the typological counting distinction of Chierchia, 1998), analyze the phenomenon with functional data (controlling for as many variables as possible), and find a high degree of uniformity in the cross-linguistic results.

With regard to the differences between the Chinese, English, and French results, some differences are observed in the regions which show an increase in activation. Only in the English results is the left ATL identified, and the French results are much more bilateral than the English and Chinese results. Curiously, though, the French results do not implicate the right cerebellum. Some possible explanations for the observed differences include differences in salience and location of number marking. While the semantic number of the non-number marked Chinese observations is conveyed to the listener through discourse cues, for the French observations, number is overtly marked on the determiner which occurs before the noun and for the English observations, number is morphologically marked on the noun itself. Another potential factor is the differences between the three datasets. They were collected by different researchers in different facilities and the Chinese and English datasets have a higher resolution than the French dataset, which leads to a more aggressive FWE correction.

4.6. Interpretation with respect to psycholinguistics

To recapitulate, Tanner et al. (2014), Tucker et al. (2015), and Wagers et al. (2009) document the plural complexity effect, which says that plural nouns are more difficult to process than singular nouns. Some of the proposed explanations for the effect include greater conceptual complexity, semantic integration, and greater morphological complexity. We discuss our results with respect to conceptual complexity and semantic integration in Section 4.2. With respect to the increased morphological complexity theory, in Arabic agreement attraction experiments, Tucker et al. (2015) find that plurals constructed via affixation take longer to read than singulars, while plurals constructed via ablaut do not. Thus, in this case, the plural complexity effect seems to apply to the plural nouns with additional morphological complexity. Recall, that in our experiments, there is on one end of the spectrum Chinese, with no overt number marking, and on the other end English, which marks plurality primarily through the plural suffix morpheme. The results for where semantically plural nouns increase activity are highly consistent between the two, though. Following this line of reasoning, if the plural complexity effects is caused by greater morphological complexity, then we are tapping into some other phenomenon, because Chinese does not morphologically mark number. What is more likely, though, is that the plural complexity effect is caused by one or a combination of latent linguistic processing aspects as discussed in Sections 4.2, 4.3 and 4.4.

The previous discussion was with respect to the plural complexity effect as it is documented in the *grammatical* stimuli of agreement attraction studies. With regard to the *ungrammatical* stimuli, where the agreement attraction errors are observed, the favored explanations are working memory based and involve cue-based retrieval. As our stimuli are grammatical (i.e., agreement attraction errors are not expected), it is difficult to relate our results. A direction for future research, though, would be to acquire functional data while participants process sentences like those in these agreement attraction studies. Given the trend towards working memory cue-based retrieval theories, we might expect to observe differences in activation in the IFG corresponding to working memory, retrieval, or reanalysis operations (Matchin, 2018; Matchin & Hickok, 2020) corresponding to those discussed by Tanner et al. (2014), Tucker et al. (2015), and Wagers et al. (2009).

5. Conclusion

In a naturalistic comprehension scenario, we find that plural nouns elicit greater activity than singular nouns across three typologically distinct human languages. This is consistent with morphological (Corbett, 2000; Greenberg, 1963), semantic (Farkas & de Swart, 2010), and psycholinguistic (Tucker et al., 2015; Wagers et al., 2009) accounts of plural markedness. While Chinese does not overtly mark bare nominals for number, French and English count nouns are overtly marked for number. Despite these differences, there is overlap between the regions in which an increase in activation is observed for plural over singular nouns: the left pars orbitalis, left pars triangularis, left dorsomedial and dorsolateral prefrontal cortex, and the left middle temporal, fusiform, and angular gyri.

These are well-known language network regions and, as such, doubt is cast on the initial hypothesis that number in language is processed similarly to domain-general number and quantity. Some of these areas have previously been implicated in the semantic processing of language (Binder et al., 2009; Graessner et al., 2021; Graves et al., 2010). Particular focus is placed on the DMPFC, and a role is discussed in which it hosts deep semantic representations after integrating lexical-semantic and syntactic representations from other portions of the language network (Frankland & Greene, 2020; Mas-Herrero & Marco-Pallarés, 2016; Murphy, 2020).

While seemingly a semantic phenomenon on the surface, semantic number appears to also play a role in syntactic processing and working memory. The PTL results are interpreted with the following reasoning: differences in feature encoding for singular and plural nouns lead to differences in functional activity in the PTL during feature-set construction. Moving from the PTL to the IFG, with the PTri providing working memory operations and space for syntactic construction, markedness/necessary encoding of a nondefault plurality feature could be expected to modulate functional activity there. With the POrb providing working memory operations and space for semantic construction, semantically plural nominals would impose a greater demand on the POrb than singular nominals through the retrieval and/or maintenance of sets of individuals as compared to a single individual. Precise formalization of all of these ideas in one, unified neurocomputational model awaits future work.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This material is based upon work supported by the National Science Foundation under Grant No. 1903783. The Chinese data collection was supported by the Jeffrey Sean Lehman Fund for Scholarly Exchange with China at Cornell University.

References

Abraham, A., Pedregosa, F., Eickenberg, M., Gervais, P., Mueller, A., Kossaifi, J., Gramfort, A., Thirion, B., & Varoquaux, G. (2014). Machine learning for neuroimaging with scikit-learn. *Frontiers in Neuroinformatics, 8*, 14. https://doi.org/10.3389/fninf.2014.00014

Ayyash, D., Malik-Moraleda, S., Gallée, J., Affourtit, J., Hoffman, M., Mineroff, Z., Jouravlev, O., & Fedorenko, E. (2021). The universal language network: A cross-linguistic investigation spanning 45 languages and 11 language families. *bioRxiv*, https://doi.org/10.1101/2021.07.28.454040.

Binder, J. R., Desai, R. H., Graves, W. W., & Conant, L. L. (2009). Where is the semantic system? A critical review and meta-analysis of 120 functional neuroimaging studies. *Cerebral Cortex, 19*, 2767–2796. https://doi.org/10.1093/cercor/bhp055

Blank, I. A., & Fedorenko, E. (2017). Domain-general brain regions do not track linguistic input as closely as language-selective regions. *Journal of Neuroscience, 37*, 9999–10011. https://doi.org/10.1523/JNEUROSCI.3642-16.2017

Bock, K., & Cutting, J. C. (1992). Regulating mental energy: Performance units in language production. *Journal of Memory and Language, 31*, 99–127. https://doi.org/10.1016/0749-596X(92)90007-K

Bock, K., & Eberhard, K. M. (1993). Meaning, sound and syntax in English number agreement. *Language and Cognitive Processes, 8*, 57–99. https://doi.org/10.1080/01690969308406949

Bock, K., & Miller, C. A. (1991). Broken agreement. *Cognitive psychology, 23*, 45–93. https://doi.org/10.1016/0010-0285(91)90003-7

Bullmore, E. T., Brammer, M. J., Rabe-Hesketh, S., Curtis, V. A., Morris, R. G., Williams, S. C., Sharma, T., & McGuire, P. K. (1999). Methods for diagnosis and treatment of stimulus-correlated motion in generic brain activation studies using fMRI. *Human Brain Mapping, 7*, 38–48. https://doi.org/10.1002/(SICI)1097-0193(1999)7:1<38::AID-HBM4>3.0.CO;2-Q

Carlson, G.N. (1977). Reference to kinds in English. Ph.D. thesis University of Massachusetts Amherst. URL https://www.proquest.com/docview/302863258?pq-origsite=gscholar&fromopenview=true.

Carreiras, M., Carr, L., Barber, H. A., & Hernandez, A. (2010). Where syntax meets math: Right intraparietal sulcus activation in response to grammatical number agreement violations. *NeuroImage, 49*, 1741–1749. https://doi.org/10.1016/j.neuroimage.2009.09.058

Castelli, F., Glaser, D. E., & Butterworth, B. (2006). Discrete and analogue quantity processing in the parietal lobe: A functional mri study. *Proceedings of the National Academy of Sciences, 103*, 4693–4698. https://doi.org/10.1073/pnas.0600444103

Chierchia, G. (1998). Reference to kinds across language. *Natural Language Semantics, 6*, 339–405. https://doi.org/10.1023/A:1008324218506

Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement, 20*, 37–46. https://doi.org/10.1177/001316446002000104

Corbett, G. G. (2000). *Number.* Cambridge University Press.

Cox, R. W. (1996). Afni: Software for analysis and visualization of functional magnetic resonance neuroimages. *Computers and Biomedical Research, 29*, 162–173. https://doi.org/10.1006/cbmr.1996.0014

Crinion, J., Turner, R., Grogan, A., Hanakawa, T., Noppeney, U., Devlin, J. T., Aso, T., Urayama, S., Fukuyama, H., Stockton, K., et al. (2006). Language control in the bilingual brain. *Science, 312*, 1537–1540. https://doi.org/10.1126/science.1127761

Dayal, V. (2011). Bare noun phrases. In K. von Heusinger, C. Maienborn, & P. Portner (Eds.), *Semantics: An International Handbook of Natural Language Meaning* (pp. 1088–1109). De Gruyter Mouton. https://doi.org/10.1515/9783110255072.1088.

Dayal, V., & Sağ, Y. (2020). Determiners and bare nouns. *Annual Review of Linguistics, 6*, 173–194. https://doi.org/10.1146/annurev-linguistics-011718-011958

Dehaene, S. (1992). Varieties of numerical abilities. *Cognition, 44*, 1–42. https://doi.org/10.1016/0010-0277(92)90049-N

Dehaene, S. (2001). Précis of the number sense. *Mind & Language, 16*, 16–36. https://doi.org/10.1111/1468-0017.00154

Dehaene, S., Piazza, M., Pinel, P., & Cohen, L. (2003). Three parietal circuits for number processing. *Cognitive Neuropsychology, 20*, 487–506. https://doi.org/10.1080/02643290244000239

Duncan, J. (2010). The multiple-demand (MD) system of the primate brain: Mental programs for intelligent behaviour. *Trends in Cognitive Sciences, 14*, 172–179. https://doi.org/10.1016/j.tics.2010.01.004

Eberhard, K. M. (1997). The marked effect of number on subject–verb agreement. *Journal of Memory and language, 36*, 147–164. https://doi.org/10.1006/jmla.1996.2484

Eberhard, K. M., Cutting, J. C., & Bock, K. (2005). Making syntax of sense: Number agreement in sentence production. *Psychological review, 112*, 531–559. https://doi.org/10.1037/0033-295X.112.3.531

Farkas, D., & de Swart, H. (2010). The semantics and pragmatics of plurals. *Semantics and Pragmatics, 3*, 1–54. https://doi.org/10.3765/sp.3.6

Fayol, M., Largy, P., & Lemaire, P. (1994). Cognitive overload and orthographic errors: When cognitive overload enhances subject–verb agreement errors. A study in French written language. *The Quarterly Journal of Experimental Psychology, 47*, 437–464. https://doi.org/10.1080/14640749408401119

Fedorenko, E., Behr, M. K., & Kanwisher, N. (2011). Functional specificity for high-level linguistic processing in the human brain. *Proceedings of the National Academy of Sciences, 108*, 16428–16433. https://doi.org/10.1073/pnas.1112937108

Fedorenko, E., Duncan, J., & Kanwisher, N. (2013). Broad domain generality in focal regions of frontal and parietal cortex. *Proceedings of the National Academy of Sciences, 110*, 16616–16621. https://doi.org/10.1073/pnas.1315235110

Flick, G., & Pylkkänen, L. (2020). Isolating syntax in natural language: MEG evidence for an early contribution of left posterior temporal cortex. *Cortex, 127*, 42–57. https://doi.org/10.1016/j.cortex.2020.01.025

Frankland, S. M., & Greene, J. D. (2020). Two ways to build a thought: Distinct forms of compositional semantic representation across brain regions. *Cerebral Cortex, 30*, 3838–3855. https://doi.org/10.1093/cercor/bhaa001

Gandour, J., Dzemidzic, M., Wong, D., Lowe, M., Tong, Y., Hsieh, L., Satthamnuwong, N., & Lurito, J. (2003). Temporal integration of speech prosody is shaped by language experience: An fMRI study. *Brain and Language, 84*, 318–336. https://doi.org/10.1016/S0093-934X(02)00505-9

Graessner, A., Zaccarella, E., & Hartwigsen, G. (2021). Differential contributions of left-hemispheric language regions to basic semantic composition. *Brain Structure and Function, 226*, 501–518. https://doi.org/10.1007/s00429-020-02196-2
Graves, W. W., Binder, J. R., Desai, R. H., Conant, L. L., & Seidenberg, M. S. (2010). Neural correlates of implicit and explicit combinatorial semantic processing. *Neuroimage, 53*, 638–646. https://doi.org/10.1016/j.neuroimage.2010.06.055
Greenberg, J. H. (1963). *Universals of language*. MIT Press.
Hale, J. T., Campanelli, L., Li, J., Bhattasali, S., Pallier, C., & Brennan, J. R. (2022). Neurocomputational models of language processing. *Annual Review of Linguistics, 8*, 427–446. https://doi.org/10.1146/annurev-linguistics-051421-020803
Hartsuiker, R. J., Schriefers, H. J., Bock, K., & Kikstra, G. M. (2003). Morphophonological influences on the construction of subject-verb agreement. *Memory & Cognition, 31*, 1316–1326. https://doi.org/10.3758/BF03195814
Honey, C. J., Thompson, C. R., Lerner, Y., & Hasson, U. (2012). Not lost in translation: Neural responses shared across languages. *Journal of Neuroscience, 32*, 15277–15283. https://doi.org/10.1523/JNEUROSCI.1800-12.2012
Humphries, C., Binder, J. R., Medler, D. A., & Liebenthal, E. (2006). Syntactic and semantic modulation of neural activity during auditory sentence comprehension. *Journal of Cognitive Neuroscience, 18*, 665–679. https://doi.org/10.1162/jocn.2006.18.4.665
Kadosh, R. C., & Walsh, V. (2009). Numerical representation in the parietal lobes: Abstract or not abstract? *Behavioral and Brain Sciences, 32*, 313–328. https://doi.org/10.1017/S0140525X09990938
Klein, D., Zatorre, R. J., Chen, J. K., Milner, B., Crane, J., Belin, P., & Bouffard, M. (2006). Bilingual brain organization: A functional magnetic resonance adaptation study. *Neuroimage, 31*, 366–375. https://doi.org/10.1016/j.neuroimage.2005.12.012
Kundu, P., Inati, S. J., Evans, J. W., Luh, W.-M., & Bandettini, P. A. (2012). Differentiating bold and non-bold signals in fMRI time series using multi-echo EPI. *Neuroimage, 60*, 1759–1770. https://doi.org/10.1016/j.neuroimage.2011.12.028
Lacadie, C., Fulbright, R., Arora, J., Constable, R., & Papademetris, X. (2008a). Brodmann areas defined in mni space using a new tracing tool in bioimage suite. In *Proceedings of the 14th annual meeting of the organization for human brain mapping*.
Lacadie, C. M., Fulbright, R. K., Rajeevan, N., Constable, R. T., & Papademetris, X. (2008b). More accurate Talairach coordinates for neuroimaging using non-linear registration. *Neuroimage, 42*, 717–725. https://doi.org/10.1016/j.neuroimage.2008.04.240
Li, J., Bhattasali, S., Zhang, S., Franzluebbers, B., Luh, W.M., Spreng, R.N., Brennan, J.R., Yang, Y., Pallier, C., & Hale, J.T. (2021). Le Petit Prince: A multilingual fMRI corpus using ecological stimuli. *bioRxiv*, https://doi.org/10.1101/2021.10.02.462875.
Li, P., Jin, Z., & Tan, L. H. (2004). Neural representations of nouns and verbs in Chinese: an fMRI study. *Neuroimage, 21*, 1533–1541. https://doi.org/10.1016/j.neuroimage.2003.10.044
Link, G. (1983). The logical analysis of plurals and mass terms: A lattice-theoretical approach. In R. Bäuerle, C. Schwarze, & A. von Stechow (Eds.), *Meaning, Use, and Interpretation of Language* (pp. 302–323). De Grutyer. https://doi.org/10.1515/9783110852820.
Lund, T. E., Madsen, K. H., Sidaros, K., Luo, W., & Nichols, T. E. (2006). Non-white noise in fMRI: Does modelling have an impact? *Neuroimage, 29*, 54–66. https://doi.org/10.1016/j.neuroimage.2005.07.005
Mas-Herrero, E., & Marco-Pallarés, J. (2016). Theta oscillations integrate functionally segregated sub-regions of the medial prefrontal cortex. *Neuroimage, 143*, 166–174. https://doi.org/10.1016/j.neuroimage.2016.08.024
Matchin, W., & Hickok, G. (2020). The cortical organization of syntax. *Cerebral Cortex, 30*, 1481–1498. https://doi.org/10.1093/cercor/bhz180
Matchin, W. G. (2018). A neuronal retuning hypothesis of sentence-specificity in Broca's area. *Psychonomic Bulletin & Review, 25*, 1682–1694. https://doi.org/10.3758/s13423-017-1377-6
Murphy, E. (2020). *The Oscillatory Nature of Language*. Cambridge University Press. https://doi.org/10.1017/9781108864466
Obleser, J., Meyer, L., & Friederici, A. D. (2011). Dynamic assignment of neural resources in auditory comprehension of complex sentences. *Neuroimage, 56*, 2310–2320. https://doi.org/10.1016/j.neuroimage.2011.03.035
Pallier, C., Devauchelle, A. D., & Dehaene, S. (2011). Cortical representation of the constituent structure of sentences. *Proceedings of the National Academy of Sciences, 108*, 2522–2527. https://doi.org/10.1073/pnas.1018711108
Paulesu, E., McCrory, E., Fazio, F., Menoncello, L., Brunswick, N., Cappa, S. F., Cotelli, M., Cossu, G., Corte, F., Lorusso, M., et al. (2000). A cultural effect on brain function. *Nature Neuroscience, 3*, 91–96. https://doi.org/10.1038/71163
Pearlmutter, N. J., Garnsey, S. M., & Bock, K. (1999). Agreement processes in sentence comprehension. *Journal of Memory and Language, 41*, 427–456. https://doi.org/10.1006/jmla.1999.2653
Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., et al. (2011). Scikit-learn: Machine learning in python. *The Journal of Machine Learning Research, 12*, 2825–2830.
Quirk, R., Greenbaum, S., Leech, G. N., & Svartvik, J. (1972). *A grammar of contemporary English*. Longman.
Rothstein, S. (2017). *Semantics for counting and measuring*. Cambridge University Press. https://doi.org/10.1017/9780511734830
Rullmann, H. (2002). Bound-variable pronouns and the semantics of number. In *Proceedings of the Western Conference on Linguistics* (pp. 243–254).
Rullmann, H., & You, A. (2006). General number and the semantics and pragmatics of indefinite bare nouns in Mandarin Chinese. In K. Turner, & K. von Heusinger (Eds.), *Where Semantics meets Pragmatics* (pp. 175–196). Elsevier.
de Saint-Exupéry, A. (1946). Le petit prince [The little prince].
Tanner, D., Nicol, J., & Brehm, L. (2014). The time-course of feature interference in agreement comprehension: Multiple mechanisms and asymmetrical attraction. *Journal of Memory and Language, 76*, 195–215. https://doi.org/10.1016/j.jml.2014.07.003
Tucker, M. A., Idrissi, A., & Almeida, D. (2015). Representing number in the real-time processing of agreement: Self-paced reading evidence from Arabic. *Frontiers in Psychology, 6*, 347. https://doi.org/10.3389/fpsyg.2015.00347
Varga, D., Halácsy, P., Kornai, A., Nagy, V., Németh, L., & Trón, V. (2005). Parallel corpora for medium density languages. In *Proceedings of RANLP 2005* (pp. 590–596). https://doi.org/10.1075/cilt.292.32var
Wagers, M. W., Lau, E. F., & Phillips, C. (2009). Agreement attraction in comprehension: Representations and processes. *Journal of Memory and Language, 61*, 206–237. https://doi.org/10.1016/j.jml.2009.04.002
Yang, R. (2001). *Common nouns, classifiers, and quantification in Chinese*. Ph.D. thesis. Rutgers University.
Zandvoort, R. (1961). *Varia syntactica. In Language and society: Essays presented to Arthur M. Jensen on his seventieth birthday* (pp. 193–203). Det Berlingske Bogtrykkeri.