Brain & Language 127 (2013) 415–427

Contents lists available at ScienceDirect

Brain & Language

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

The representation of category typicality in the frontal cortex and its cross-linguistic variations

Chao Liu [a], Twila Tardif [b,c,*], Haiyan Wu [a], Christopher S. Monk [b,c], Yue-Jia Luo [a], Xiaoqin Mai [d,*]

[a] State Key Laboratory of Cognitive Neuroscience and Learning, Beijing Normal University, Beijing 100875, China
[b] Department of Psychology, University of Michigan, Ann Arbor, MI 48109-1109, USA
[c] Center for Human Growth and Development, University of Michigan, Ann Arbor, MI 48109-5406, USA
[d] Department of Psychology, Renmin University of China, Beijing 100872, China

ARTICLE INFO

Article history:
Accepted 3 September 2013
Available online 14 October 2013

Keywords:
Frontal cortex
Categorization
Typicality effect
fMRI
Cross-linguistic

ABSTRACT

When asked to judge the membership of typical (e.g., car) vs. atypical (e.g., train) pictures of a category (e.g., vehicle), native English ($N = 18$) and native Chinese speakers ($N = 18$) showed distinctive patterns of brain activity despite showing similar behavioral responses. Moreover, these differences were mainly due to the *amount* and *pervasiveness* of category information linguistically embedded in the everyday names of the items in the respective languages, with important differences across languages in how pervasive category labels are embedded in item-level terms. Nonetheless, the left inferior frontal gyrus and the bilateral medial frontal gyrus are the most consistent neural correlates of category typicality that persist across languages and linguistic cues. These data together suggest that both cross- and within-language differences in the explicitness of category information have strong effects on the nature of categorization processes performed by the brain.

1. Introduction

One of the fundamental insights into semantic memory is the role of typicality in both structuring and providing access to members of a category (Mervis, Catlin, & Rosch, 1976). When asked questions such as "*Is an ostrich a bird?*" or "*Is a robin a bird?*," the general behavioral finding (Mervis & Rosch, 1981) is that people respond more quickly and more accurately to "robin" than "ostrich" (i.e., demonstrate a "typicality effect") simply because "robin" is a more *typical* example of the category "bird" than "ostrich."

As one of the most consistent indexes of categorization processes in behavioral studies (Mervis & Rosch, 1981), the typicality effect has also been investigated with neuroimaging techniques such as the Event-Related Potential (ERP). ERP studies have found that typicality effects in linguistic stimuli are marked by a N400 component, such that atypical items of a category elicit a larger N400 than typical items, regardless of the frequency of the item labels (Fujihara, Nageishi, Koyama, & Nakajima, 1998; Heinze, Muente, & Kutas, 1998; Stuss, Picton, & Cerri, 1988). In addition to the N400 component found in the frontal, temporal, and parietal areas of the brain, studies with pictorial stimuli have found additional components at 160 ms (P160) in occipital areas and 280–300 ms (N300) in other posterior areas, representing additional perceptual and semantic processing of pictorial atypical vs. typical stimuli (Barrett & Rugg, 1990; Hauk et al., 2007; McPherson & Holcomb, 1999).

However, to the best of our knowledge, this classic "typicality effect" has not yet been directly investigated using neuroimaging techniques with high spatial resolution such as functional Magnet Reasoning Imaging (fMRI) (Patterson, 2007). Nonetheless, there have been some studies that have begun to shed light on what one might expect for such an effect. Studies investigating the orthographic typicality (e.g., CHEESE is a typical English word but SEIZE is an atypical one) or phonetic typicality (e.g., sounds belong to normal human voicing continuum or not) of word stimuli, for example, showed that atypical items elicited greater activation than typical items in language processing areas such as the left inferior frontal region and bilateral superior temporal regions (Myers, 2007; Woollams, Silani, Okada, Patterson, & Price, 2011). Moreover, fMRI studies using other categorization processing paradigms with English-speaking adults including both healthy controls and patients with semantic dementia have identified three qualitatively different categorization systems in the brain (Smith & Grossman, 2008). The first is a rule-based categorization process associated with a working memory system and selective attention in the frontal and parietal areas, especially the left inferior frontal gyrus. The second is a similarity-based categorization associated with explicit long-term memory and integration of perceptual

---

* Corresponding authors. Address: Department of Psychology, University of Michigan, Ann Arbor, MI 48109-1109, USA. Fax: +86 734 936 9288 (T. Tardif), fax: +86 10 82506321 (X. Mai).

E-mail addresses: Twila@umich.edu (T. Tardif), maixq@ruc.edu.cn (X. Mai).

0093-934X/$ - see front matter © 2013 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.bandl.2013.09.002

features in the parietal-temporal areas. For the third, other sorts of implicit categorization processes associated with implicit long term memory in the temporal-occipital areas have also been identified. However, which of these three categorization systems might be involved in the typicality effect is still unknown. This question is particularly interesting for semantic typicality processing. Compared with other typicality processing such as orthographic typicality or phonetic typicality, semantic typicality processing is characterized by a more complicated connectionist representation where concepts correspond to distributed representations occupying positions in a multidimensional semantic space (Patterson, 2007). Studies on the brain mechanism of related semantic processing have found that categorization of both word and pictorial stimuli show activation in the bilateral middle and inferior frontal gyrus, the bilateral inferior parietal lobule, bilateral temporal-occipital conjunctions, anterior cingulate cortex and caudate (Adams & Janata, 2002; Ganis, Schendan, & Kosslyn, 2007; Grossman et al., 2002; Jiang et al., 2007; Koenig et al., 2005; Myers, 2007; Reber, Gitelman, Parrish, & Mesulam, 2003). The left inferior frontal gyrus, in particular, has been identified as a region which contributes greatly to semantic and lexical access (Bookheimer, 2002; Hagoort, Hald, Bastiaansen, & Petersson, 2004; Vandenberghe, Price, Wise, Josephs, & Frackowiak, 1996). Thus, our first aim of the present study is to investigate the neural correlates of the semantic typicality effect by recording an MRI signal when participants perform a classic category verification task with pictorial stimuli. We would expect to identify a brain network involved in semantic processing, particularly in the inferior frontal cortex.

The second aim of the present study is to investigate how the neural correlates of the typicality effect might be similar or differ across languages. The idea that language plays an important role in categorization is not new in cognitive psychology. As Whorf suggested in the linguistic relativity hypothesis, "We dissect nature along lines laid down by our native language."(Whorf, 1956). Developmental studies have shown that language shapes the way that object categories are organized and structured in children's minds (Martinez & Shatz, 1996; Yoshida & Smith, 2003). In addition, cross-linguistic studies have found that different languages differ greatly in providing linguistic cues to a word's semantic category. For example, in English, basic level object nouns usually do not share any obvious relationships to their superordinate category labels (e.g., nouns for wheeled vehicles are bicycle, truck, car, taxi, bus, train, etc.), although some do (e.g., cuttlefish, catfish). In contrast, most basic level object nouns in Mandarin Chinese contain superordinate category information in some way, either by sharing a common root morpheme (Tardif, 2006) (e.g., all wheeled vehicles share the common morpheme *che1* (车) that means "vehicle", such as bicycle – *zi4xing2che1*自行车, truck – *ka3che1*卡车, car – *jiao4che1*轿车, taxi – *chu1zu1che1*出租车, bus – *gong1gong4qi4che1*公共汽车, train – *huo3che1*火车), or by including a unpronounceable orthographic "radical" that cues either the basic or superordinate category in the written character (Zhou, Marslen-Wilson, Taft, & Shu, 1999; Zhou, 1978) (e.g., the noun "fish" (*yu2*鱼)) is not only a simple character in its own right, but is also an orthographic component (also known as a "semantic radical") in the written character for different fish names, such as carp – *li3*鲤, bass – *lu2*鲈, catfish – *nian2*鲇, and shark – *sha*鲨). Over 80% of Chinese characters provide semantic radicals (Zhou et al., 1999; Zhou, 1978), and these can be traced back to the oracle bone characters used 3500 years ago (e.g., the radical of "water" *shui3*水 in the characters of river – *he2* 河 and wine – *jiu3* 酒), thus creating a fascinating and long-standing tradition of cueing category information that is pervasive in basic level terms in Chinese.

In summary, in English, nouns tend to have opaque or "non-transparent" cues to categories, whereas Chinese nouns have a highly productive and pervasive morphological and orthographic compounding system which provides explicit cues to category membership. How could these language differences then influence categorization processes and the typicality effect? Since both typical and atypical nouns in Chinese contain exactly the same linguistic cue (whether morphological or orthographic), it is possible that this pervasive system of cues might be used to aid Chinese speakers in making category judgments and thus obviate the need for typicality as a cue. This hypothesis was supported in a series of cross-cultural ERP studies comparing English and Chinese speaking adults with a category verification task. In these studies, Chinese speaking adults showed no N300 or N400 components revealing no apparent differences in the processing of typical vs. atypical items (Liu et al., 2010) despite clear differences with the identical stimuli for English-speaking adults and strong similarities in the N300 and N400 effects shown across languages for within vs. out of category items. These studies led to the suggestion that the linguistic cues in Chinese nouns facilitated the semantic access of category information in Chinese speaking adults and eliminated the left frontal N300 and N400 typicality effects. However, the locus of these cross-linguistic similarities and differences are still not clear.

In the current study, we investigated the neural correlates of the typicality effect and its cross-linguistic variations with native US English and native Chinese speakers using event-related fMRI. We conducted a category verification task with pictorial stimuli that differ in both typicality and linguistic cues to category membership and demonstrate that these linguistic cues are responsible for differing levels of brain activation in the same brain areas identified by others as responsible for semantic categorization, despite overall similarities in behavioral responses.

2. Methods

2.1. Participants

Twenty native Mandarin Chinese speakers and 19 native US English speakers (from US or Canada) in Beijing, all right-handed with normal vision, participated in this study and were each paid RMB100 (approximately US$12). Considering previous studies have demonstrated differences between bilinguals and monolinguals in naming task (Ameel, Malt, Storms, & Van Assche, 2009; Ameel, Storms, Malt, & Sloman, 2005), we also controlled the Language 2 (L2) level in participants recruiting. To minimize the proficiency of English, all native Chinese speakers were college students who had not yet passed College English Test (CET). To minimize the proficiency of Chinese for native English speakers, they were required to have lived in Beijing for less than three years and have learned Chinese for less than one year. Three participants were excluded from further analysis, two Chinese speakers for poor behavioral performance and one English speaker for uncorrectable head movement (>4 mm) during fMRI acquisition. The final sample consisted of 18 Chinese speakers (10 females, *M* age = 22.33 years) and 18 English speakers (10 females, *M* age = 25.38 years) in the behavioral and fMRI data analysis. The recruitment of participants in Beijing was approved by IRBs at Beijing Normal University and the University of Michigan (B04-00001580-M1).

2.2. Stimuli

Twenty-eight grayscale object pictures of 14 categories were selected from previous ERP studies (Liu et al., 2010) (Fig. 1), for which two pilot studies were conducted to refine and ensure the cross-linguistic comparability of the pictorial stimuli and their judged typicality. In Pilot Study one, 25 English and 25 Chinese

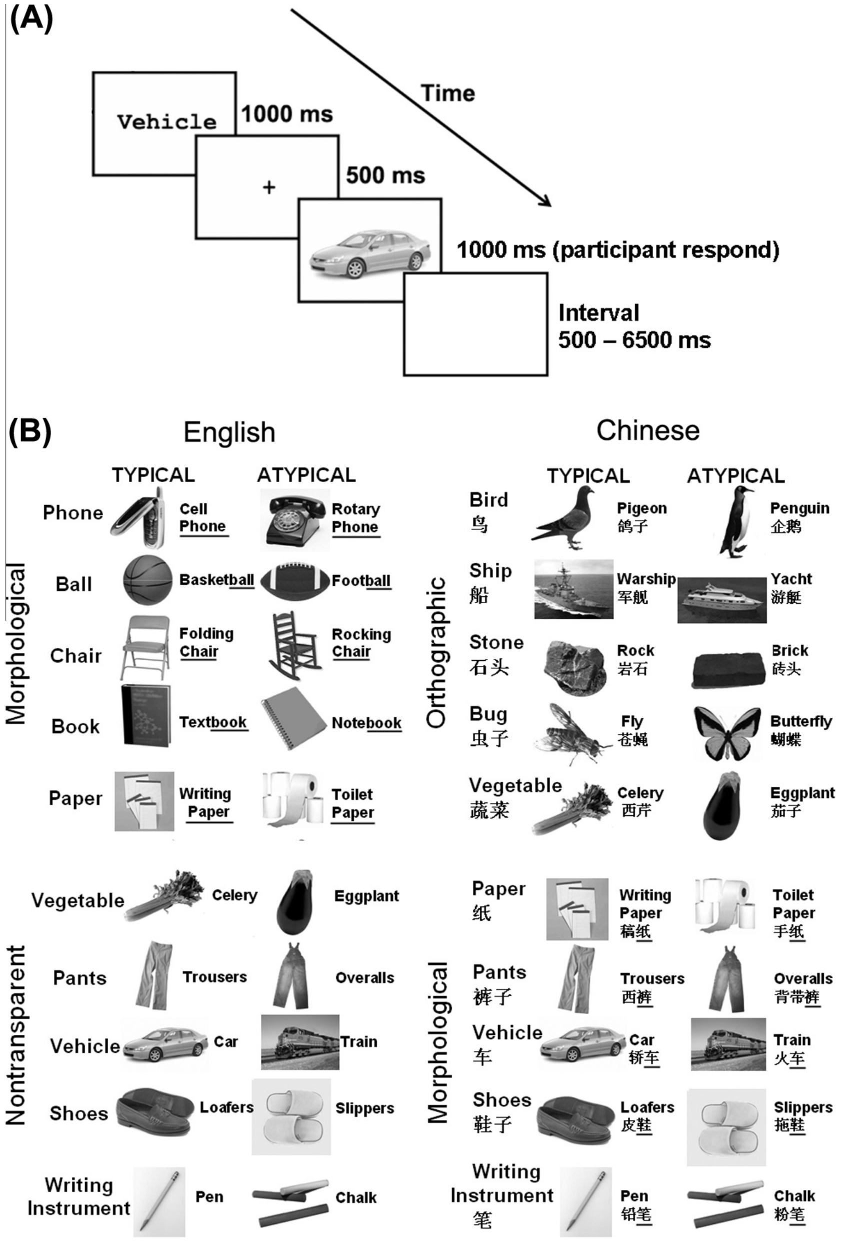

Fig. 1. Experimental procedure (A) and materials (B). Participants first saw a category-level label (e.g., "VEHICLE" in English or che1 "车" in Chinese), followed by a picture of either a typical, atypical or out-of-category object (e.g., a car, a train or a pen, respectively). The participants' task was to judge whether the object picture was an exemplar of the label or not. In English, 5 of the ten categories included nontransparent items only (e.g., VEHICLE: car, train) and 5 included morphologically transparent items (e.g., PAPER: writing paper, tissue paper). In Chinese, 5 of the ten categories included morphologically transparent items (e.g., VEHICLE/che1车: car -jiao4che1轿车, train – huo3che1火车), and 5 were orthographically transparent (e.g., BUG/chong2虫, fly -cang1ying苍蝇, butterfly – hu2die2蝴蝶). Among these ten English and ten Chinese categories, six included identical categories and item pictures across languages (bottom six categories in both languages, 5 nontransparent and 1 morphological for English, 5 morphological and 1 orthographic for Chinese). In addition, all participants also completed a typicality rating survey in a 1–6 scale before the scan (see Method).

participants were given a questionnaire about the acceptability of replacing certain terms for each other (e.g., "Can car be used to replace the word vehicle?"). For example, given the nouns fly cang1ying1苍蝇, worm qiu1ying3 蚯蚓, bug chong2zi 虫, and mosquito wen2zi 蚊子(from the category BUG), a participant could say that "fly" can be used in place of "bug", or that "fly" can be used in place of "mosquito", and so on. In Pilot Study two, 29 English and 24 Chinese participants were asked to rate the typicality of each picture, given either the category- (e.g., vehicle) or item- (e.g., car) level label, on a six-point scale. These ratings were then used to identify typical and atypical items for each category. Additionally, to ensure consistency across categories and languages, every participant completed a typicality rating survey before the scan by first naming the item stimulus picture and then rating the typicality of each item on a 1–6 point scale, with 1 representing *not at all typical* (完全不典型) and 6 representing *extremely typical* (极端典型), in response to the visually presented question "How typical is this as an example of a VEHICLE?" (这在多大程度上是一台典型的车). Any responses that did not contain the expected orthographic/ morphological parts in the naming task were counted as inaccurate labels. All stimulus pictures received >80% naming accuracy at the item-level from both English (86.3%) and Chinese (84.4%) speakers. The typicality rating results for the six categories compared cross-linguistically and ten categories in English and Chinese revealed no significant interactions between typicality and other factors of interest such as language or label type

($Ps > 0.46$), thus ensuring the comparability of the stimuli across languages and conditions in the current study.

2.3. Procedure and task

A schematic summary of the procedure and task can be found in Fig. 1. Specifically, there were two types of categories (label types) for each language: those with morphologically transparent vs. nontransparent cues to the superordinate category for English and morphologically transparent vs. orthographically transparent cues for Chinese. Within these two types of categories, one pictorial exemplar each of a "typical" (e.g., car) and "atypical" (e.g., train) category member was shown to participants, who were asked to judge simply whether the pictured item "was" (yes) or "was not" (no) a member of the category whose term appeared preceding the presentation of the pictorial stimuli.

All participants finished 12 practice trials with category and item pictures that were not selected for the real experiment (e.g., category "station" *che1zhan4* 车站 and a picture of a train station) before the scan. During the fMRI scan, a total of 400 trials, 20 trials for each of the 20 pictures, were presented in random order to each participant in four separate runs with 100 trials each. The intertrial interval was jittered at 500, 2000, 3500, 5000, and 6500 ms with differing probabilities (50%, 25%, 12%, 7%, 6%, respectively). Half of all trials required a *Yes* response (e.g., label "vehicle", followed by a picture of car) and half required a *No* response (e.g., label "vehicle", followed by a picture of eggplant). Among the 200 "*Yes*" trials, there were 50 trials for each condition in a fully crossed design of Typicality (Typical *vs.* Atypical) by Label type (Morphological *vs.* Nontransparent in English and Morphological *vs.* Orthographic in Chinese). The cross-linguistic comparison was done by comparing the six identical categories (vegetable, paper, pants, vehicle, shoes, writing instrument) in English and Chinese, yielding 60 trials for each language. The whole experimental session lasted approximately 1 h with 25–30 min for fMRI data acquisition.

2.4. Image acquisition

Echo Planar Imaging was acquired from a Siemens 3T scanner (TR = 1500 ms, TE = 28 ms, interleaved, 28 axial slices with 4.8-mm-thick each, field of view 200 × 200 mm, acquisition matrix was 64 × 64, flip angle 75°, in-plane resolution = 3.1 × 3.1 mm^2). A total of 1184 scans were acquired in four runs. High-resolution T1-weighted images were obtained for each subject to provide detailed anatomy (1.0 × 1.0 × 1.3).

2.5. Imaging data analysis

Data analysis was performed with SPM5 from the Wellcome Department of Cognitive Neurology, London. MNI coordinates (Friston et al., 1995) were transferred into Talairach coordinates (Talairach & Tournoux, 1988) according to the criteria specified by http://www.mrc-cbu.cam.ac.uk/Imaging/Common/mnispace.shtml. Image data were represented using MRIcroN http://www.sph.sc.edu/comd/rorden/mricron/ and CARET http://brainvis.wustl.edu/wiki/index.php/Main_Page. Talairach coordinates were transferred to brain regions using the Talairach Daemon database (Lancaster et al., 1997). The first two scans of each run were discarded from the analysis to eliminate non-equilibrium effects of magnetization. Scans were first preprocessed for slice-timing, realignment, normalization (to MNI space), and smoothing (8 × 8 × 8 mm, Gaussian spatial filter). The resulting images had voxel size of 3.13 × 3.13 × 4.8 mm^3.

Two individual-level analyses for each participant were performed separately for the cross- and within-language comparisons. For the cross-linguistic comparison, each of the eight types of trial generated from crossing the Typicality (typical *vs.* atypical) by Language (English *vs.* Chinese) by Response (Yes *vs.* No) conditions, was contrasted with the six motion parameters obtained from realignment as the covariate. For the within-language comparison, everything else was the same except that Language was replaced by the Label type (nontransparent *vs.* morphological in English and orthographic *vs.* morphological in Chinese). Long-term signal variations were eliminated with a high-pass filter set at 128 s, and a low-pass filter was achieved by convolution with the standard SPM hemodynamic response function (HRF). The hemodynamic response of the event was time-locked to the presentation of the object pictures. The duration was set to 0 as an event-related design. We performed two group-level random effects analyses. A two-sample *t*-test between English and Chinese speakers was conducted for the cross-linguistic comparison and a one-sample *t*-test among English speakers and Chinese speakers was conducted separately for the within-subjects comparison between different label types in each language. The cross-linguistic contrasts with two sample *t*-tests used an uncorrected voxelwise threshold of $P < 0.001$, $P < 0.05$ ($K > 10$) (Canli et al., 2005) with False Discovery Rate (FDR) corrected for multiple comparisons using the small volume correction (SVC), whereas the within-linguistic contrasts with one sample *t*-tests revealed weaker brain activation in general, thus a reduced threshold of uncorrected voxelwise $P < 0.005$ was set with a FDR corrected $P < 0.05$ ($K > 20$ voxels) (Depue, Curran, & Banich, 2007; Seymour, Daw, Dayan, Singer, & Dolan, 2007) with SVC correction for all conditions but one. In the Chinese morphological (Atypical vs. Typical) contrast, we expected *not* to find any differences, and thus to be more conservative in our conclusions, we set a higher threshold of FDR corrected $P < 0.1$, ($K > 10$ voxels) with SVC correction in order to reveal any potential activation that could reflect the behavioral differences in this condition. A less stringent threshold was used for the one sample t-test of the within-language comparison than for the two sample *t*-test of the between-language comparison due to the fact that we had fewer categories per condition in the within-language comparison (5 categories and 50 trials) than that in the between-language comparison (6 categories and 60 trials). Based on the frontal and parietal regions identified in various previous categorization studies (Adams & Janata, 2002; Ganis et al., 2007; Grossman et al., 2002; Jiang et al., 2007; Koenig et al., 2005; Myers, 2007; Reber et al., 2003), small-volume correction (SVC) was done separately using ROIs within the left BA 46 and 47, the right BA46 and 47, the bilateral BA 8, bilateral BA 9 and bilateral BA 40, defined by the Talairach Daemon Brodmann Areas from the WFU_PickAtlas 2.40 (Maldjian, Laurienti, Kraft, & Burdette, 2003). Average signals in the ROIs were extracted and plotted using Marsbar (Fig. 5A) (Brett, Anton, Valabregue, & Poline, 2002). The SVC in the Chinese morphological (Atypical vs. Typical) contrast was set in the left BA19 and right caudate tail. Determination of common regions for "English Morphological Atypical" and "Chinese Orthographic Atypical" was calculated by a conjunction analysis implemented in SPM5 for the contrasts "English Morphological Atypical > Typical" and "Chinese Orthographic Atypical > Typical". The threshold of the resulting statistical map was $P < 0.005$, uncorrected for at least 20 contiguous voxels (Fig. 5B).

3. Results

3.1. Behavioral results

Trials with a response time >1200 ms or <200 ms were excluded as outliers in both behavioral and fMRI analyses (English: 1.0% of all responses, Chinese: 4.5% of all responses). Only trials with correct responses were included in the behavioral and fMRI analysis.

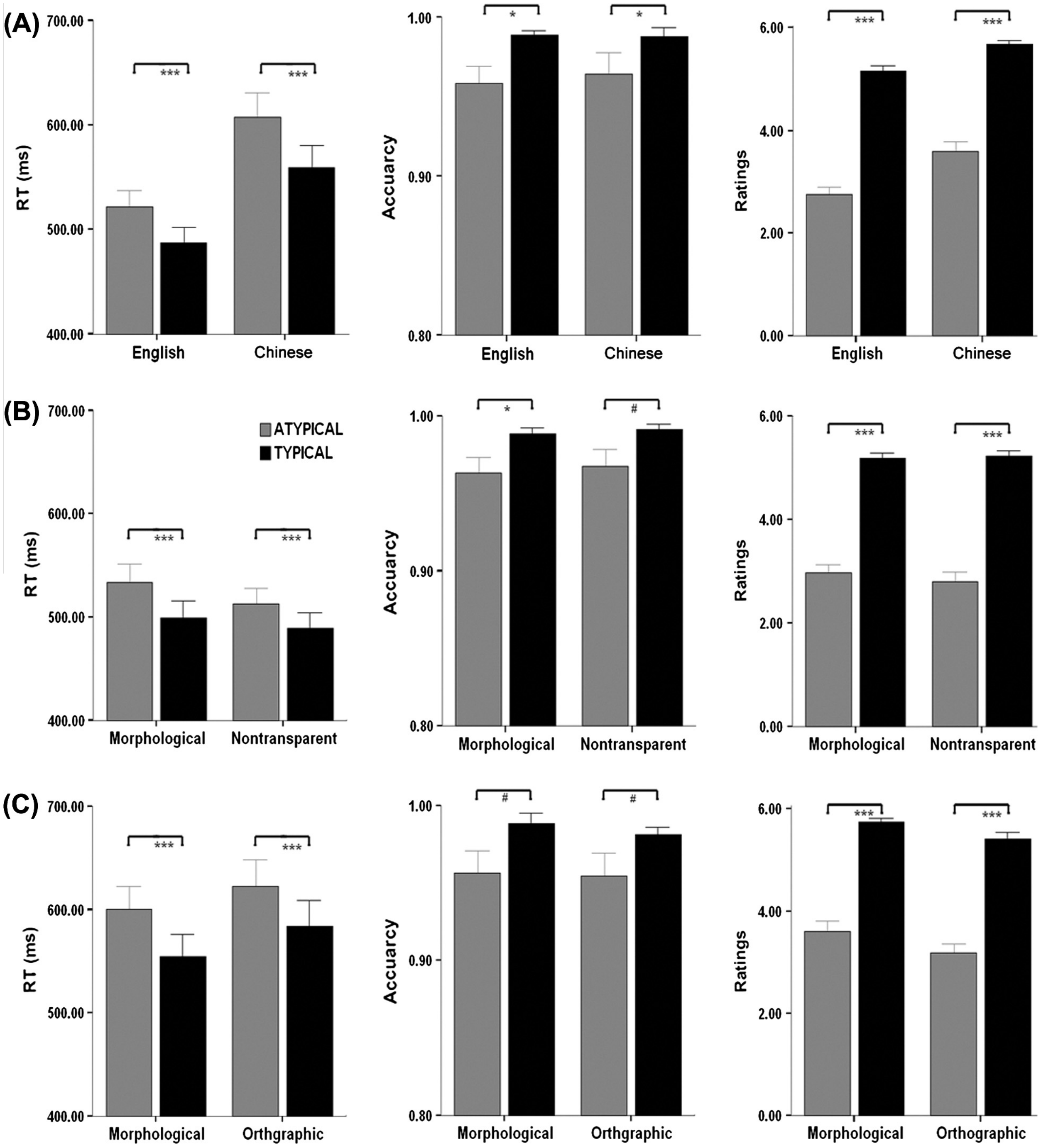

Fig. 2. Behavioral and typicality rating (1–6) results show cross-linguistic differences in the six identical categories shown for English and Chinese speakers (A), between morphologically transparent and nontransparent items for English speakers (B) and between morphologically transparent and orthographically transparent items for Chinese speakers (C). Reaction time data is presented only for correct responses. $^{\#}P < 0.01$, $^{*}P < 0.05$, $^{**}P < 0.01$ and $^{***}P < 0.001$.

A series of Typicality (Typical vs. Atypical) by Language (English vs. Chinese) repeated measures ANOVAs with Bonferroni corrections for post-hoc analyses were conducted for the accuracy, RT and rating data of all stimuli to explore the effect of Typicality and Language in the "yes" responses (Fig. 2A). We found a significant main effect of Typicality in all three measures, such that participants rated the items we included as Typical to be significantly more "typical" than those considered to be "Atypical," $F(1,34) = 308.56$, $P < 0.001$, and made more errors and responded more slowly for Atypical items than Typical items [Accuracy: $M = 0.95$ and 0.99, $F(1,34) = 10.57$, $P = 0.003$; RT: $M = 565.45$ and 523.56, $F(1,34) = 52.06$, $P < 0.001$, respectively], as shown in Fig. 2A, which repeated the classic Typicality Effect found in many previous behavioral and ERP studies (Fujihara et al., 1998; Heinze et al., 1998; Mervis & Rosch, 1981). Moreover, there was a significant main effect of Language in the RT data such that Chinese participants responded more slowly than English speakers ($M = 582.17$ and 505.34, $F(1,34) = 9.07$, $P = 0.004$) and generally gave higher ratings (Typical $M = 5.65$, Atypical $M = 3.57$) than English speakers (Typical $M = 5.13$, Atypical $M = 2.76$) for both types of items, with no significant Typicality by Language interaction in the ratings, reaction times or accuracy data. Overall, even considering the

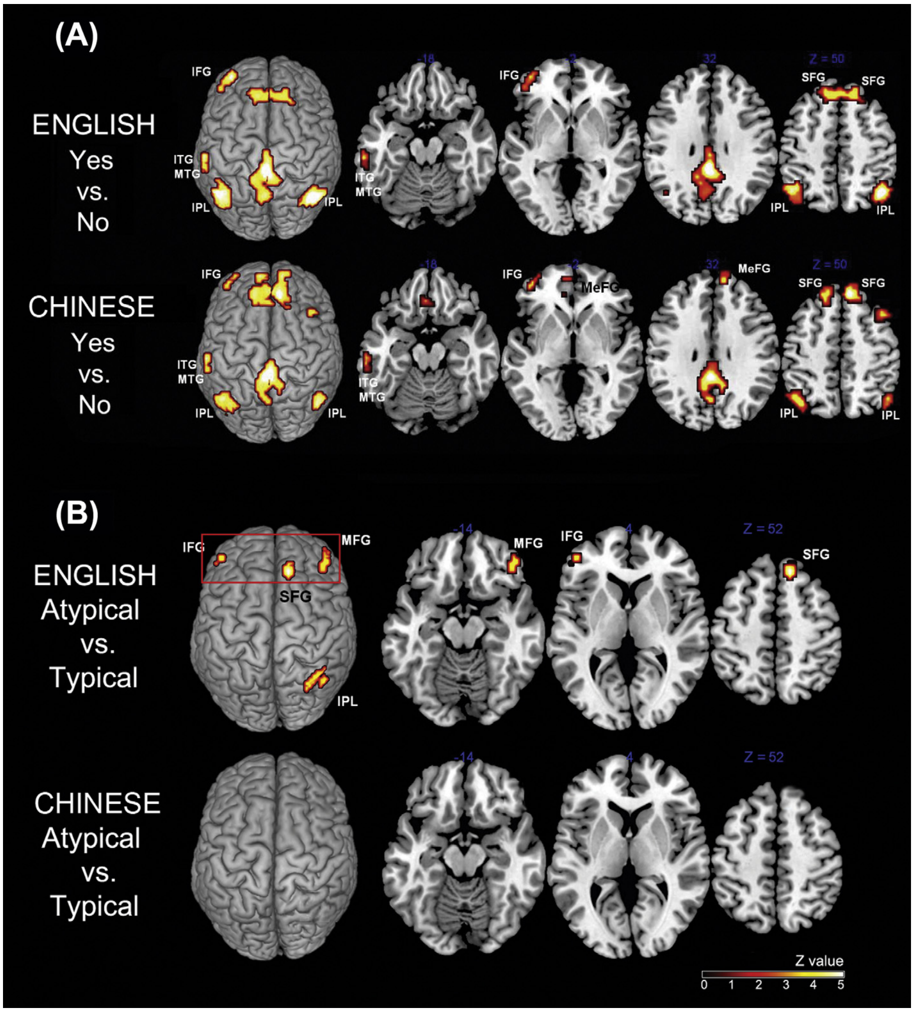

Fig. 3. Similarities in brain activation for pictures of items in all categories that received Yes responses (in category) vs. No responses (out of category) from English- and Chinese-speaking participants (A) (Table 1). Brain regions showing activation (atypical-typical) for pictures of items in six identical categories judged by English- and Chinese-speaking participants (Yes responses only) (B) (Table 2). Slices begin with the overall axial view with infinite search depth. English and Chinese participants showed almost identical brain patterns for the Yes vs. No contrast. However, the typicality effect was associated with the left IFG (BA 46), the right MFG (BA11) and the right SFG (BA8) only in English speakers when comparing those six identical categories across languages.

apparently stronger typicality effect in the RTs for the Chinese participants in this study, these results are consistent with previous studies finding typicality effects for English and other languages such as Japanese and German indicating that *behavioral* manifestations of the typicality effect are robust across languages. The following analyses thus consider, first, the effects of Label Type within each language, and, second, the neuroimaging results.

For the English speakers, a Label Type (Morphological *vs.* Nontransparent) by Typicality (Typical *vs.* Atypical) repeated measures ANOVA with Bonferroni corrections for post-hoc analyses on the accuracy, RT and rating data revealed significant typicality effects for both non-transparent and morphologically transparent labels (Fig. 2B). As with the overall ANOVAs, there were main effects of Typicality in the ratings, $F(1,17) = 271.99$, $P < 0.001$, as well as in the RT and accuracy data such that participants made more errors and responded more slowly for Atypical than Typical items (Accuracy*:* $M = 0.96$ and 0.99, $F(1,17) = 8.28$, $P = 0.010$; RT: $M = 520.86$ ms and 491.78 ms, $F(1,17) = 29.96$, $P < 0.001$). Interestingly, there was also a significant main effect of Label Type in the RT data such that participants responded more slowly for pictures with Morphologically transparent labels than pictures of items with Nontransparent labels ($M = 512.30$ ms and 500.34 ms, $F(1,17) = 8.78$, $P = 0.009$). One possible reason is that there are less Morphological labels than Nontransparent labels in English, thus such a slow response for morphological labels might be attributed to the novelty effect. However, no interactions between Typicality and Label Type were found for any of the measures and English speakers generally gave similar ratings for pictures of items Morphological cues to category membership (Typical $M = 5.19$, Atypical $M = 2.97$) as they did for Nontransparent items (Typical $M = 5.24$, Atypical $M = 2.81$).

For the Chinese speakers, a Label Type (Morphological *vs.* Orthographic) by Typicality (Typical *vs.* Atypical) repeated measures ANOVAs with Bonferroni corrections for post-hoc analyses revealed significant typicality effects for the ratings, $F(1,17) = 172.32$, $P < 0.001$, as well as for the RT and accuracy data such that participants made more errors and responded more slowly for Atypical than Typical items (Accuracy*:* $M = 0.96$ and 0.99, $F(1,17) = 5.01$, $P = 0.039$; RT: $M = 615.22$ ms and 571.51 ms, $F(1,17) = 42.11$, $P < 0.001$). In addition, there was also a significant main effect of Label Type in both the ratings, $F(1,17) = 25.77$, $P < 0.001$, and the RT data such that participants generally gave higher ratings for pictorial stimuli with Morphological (Typical $M = 5.76$, Atypical $M = 3.62$) than Orthographic cues to category membership (Typical $M = 5.42$, Atypical $M = 3.19$), and responded more slowly for items with Orthographic than Morphological

Table 1

Brain regions showing significant activation between Yes and No responses in English and Chinese. (Two sample *t*-test, uncorrected voxelwise threshold of $P < 0.001$, with $P < 0.05$ with False Discovery Rate (FDR) corrected for multiple comparisons ($K > 10$ voxels).)

Contrast	BA	*P (FDR) corrected*	Voxel	*x*	*y*	*z*	*Z*
English (Yes–No)							
L Superior Frontal G	8	<0.01	100	−13	35	47	3.92
R Superior Frontal G	8	0.02		13	35	47	3.78
L Middle Frontal G	10	0.01	35	−34	52	7	4.45
L Inferior Frontal G	10	0.03		−47	43	−2	3.50
L Inferior Parietal L	40	<0.01	105	−49	−55	47	4.82
L Inferior Temporal G	20	0.01	17	−62	−25	−19	4.31
L Middle Temporal G	21	0.01		−62	−34	−10	4.02
L Cingulate G	31	<0.01	149	0	−38	29	4.99
L Precuneus	7	0.01		−3	−59	38	4.03
R Inferior Parietal L	40	<0.01	97	47	−62	47	5.01
Chinese (Yes–No)							
L Medial Frontal G	10	0.01	18	−9	57	2	4.29
L Medial Frontal G	10	0.03	16	−9	39	−10	3.64
L Middle Frontal G	10	0.02	17	−34	57	2	3.89
L Inferior Frontal G	10	0.05		−44	46	−2	3.31
L Inferior Parietal L	40	0.02	81	−47	−59	38	4.14
L Superior Parietal L	7	0.02		−38	−64	52	3.93
L Inferior Parietal L	39	0.05		−47	−68	43	3.34
L Inferior Temporal G	20	0.01	16	−62	−22	−15	4.47
L Middle Temporal G	21	0.02		−62	−34	−10	3.82
L Cingulate G	31	<0.01	133	−3	−38	33	4.90
L Precuneus	7	0.03		−3	−59	34	3.53
R Superior Frontal G	8	<0.01	147	9	42	47	4.90
R Superior Frontal G	9	0.01		13	58	24	4.19
L Medial Frontal G	8	0.02		0	45	37	4.08
L Superior Frontal G	8	0.02		−16	45	42	3.97
R Middle Frontal G	8	0.02	12	41	24	48	3.84
R Inferior Parietal L	40	0.01	38	47	−65	47	4.34

labels ($M = 603.23$ ms and 583.49 ms, $F(1,17) = 10.25$, $P = 0.005$). As with English, no interaction was found between Typicality and Label Type for any of the measures, as can be seen from Fig. 2 C. Again, because there are less Orthographical labels than Morphological labels in Chinese, thus such a slow response for Orthographical labels might be attributed to the novelty effect.

3.2. Imaging results

When contrasting responses to items belonging to the category (*Yes* responses) versus those not belonging (*No* responses) (Fig. 3 A, Table 1), both groups of speakers showed similar patterns of activity. Specifically, out-of-category items elicited greater activity in the bilateral SFG (BA8), the left MFG and IFG (BA10), the bilateral IPL (BA40), the left inferior temporal lobe (ITL) (BA20), and the left middle temporal lobule (MTL) (BA21) (Fig. 3 A, Table 1). This indicates that Chinese and English speakers performed the task similarly when distinguishing between in-category and out-of-category items, not only at the level of categorization processing, but also at the level of basic visual processing and decision making.

In contrast, the fMRI results showed dramatic differences in the English and Chinese speakers' processing of typical *vs.* atypical items. For English speakers, atypical items elicited larger activity in the left inferior frontal gyrus (IFG) (Brodmann areas [BA] 46, 47), the right middle frontal gyrus (MFG) (BA10, 11), the right superior frontal gyrus (SFG) (BA8) and the right inferior parietal lobule (IPL) (BA40) (Fig. 3 B, Table 2). In contrast, Chinese speakers showed no differences between typical and atypical items in these regions (Fig. 3B, Table 2).

The highly similar brain pattern we found in the *Yes vs. No* contrast between Chinese and English speakers indicates that the differences in the non-overlapping items is not likely to be the primary source of cross-linguistic differences in brain activation for the *Typical vs. Atypical* comparison. Instead, we propose that the cross-linguistic difference in the typical *vs.* atypical contrast was not due to item differences or context effects, but to other factors such as the prevalence and availability of linguistically transparent cues to category membership.

Table 2

Brain regions showing significant activations between typical and atypical items in English and Chinese. (Two sample *t*-test, uncorrected voxelwise threshold of $P < 0.001$, $P < 0.05$ with False Discovery Rate (FDR) corrected for multiple comparisons with the small volume correction (SVC) in the regions of frontal and parietal cortex. ($K > 10$ voxels)).

Contrast	BA	*P(FDR)*	Voxel	*x*	*y*	*Z*	*Z*
English (Atypical–Typical)							
L Inferior Frontal G	46	0.02	13	−47	46	2	3.75
L Middle Frontal G	47	0.03		−50	40	−6	3.22
R Superior Frontal G	8	<0.01	18	13	35	47	4.51
R Middle Frontal G	11	0.05	16	44	39	−14	3.40
R Middle Frontal G	10	0.05		47	48	−7	3.37
R Inferior Parietal L	40	0.05	24	44	−49	42	3.56
Chinese (Atypical–Typical)							
(none)							

The comparison among different types of linguistic labels within English and Chinese revealed even more details about the different patterns of brain activity involved in the typicality effect. Although the behavioral results did not show interactions between the different types of linguistic labels and the typicality of the pictures for either English or Chinese (Fig. 2), brain activity to these different types of items was strikingly distinct, both within and across languages. For English speakers, semantically "nontransparent" items (e.g., VEHICLE: car) activated several distinct areas including the bilateral SFG (BA8), IFG (BA46, 10), MFG (BA8, 11) and the right IPL (BA40). In contrast, "morphologically transparent" items (e.g., BALL: basketball) activated only the bilateral medial frontal gyrus (MeFG) (BA8), the left IFG (BA47, 45), IPL (BA40) and the right lingual gyrus (BA18). Even more interestingly, Chinese "morphologically transparent" items (e.g., VEHICLE *che1*车*:* car *jiao4che1*轿车) showed only slight activation in the left middle occipital gyrus (MOG) and the right caudate without any activation

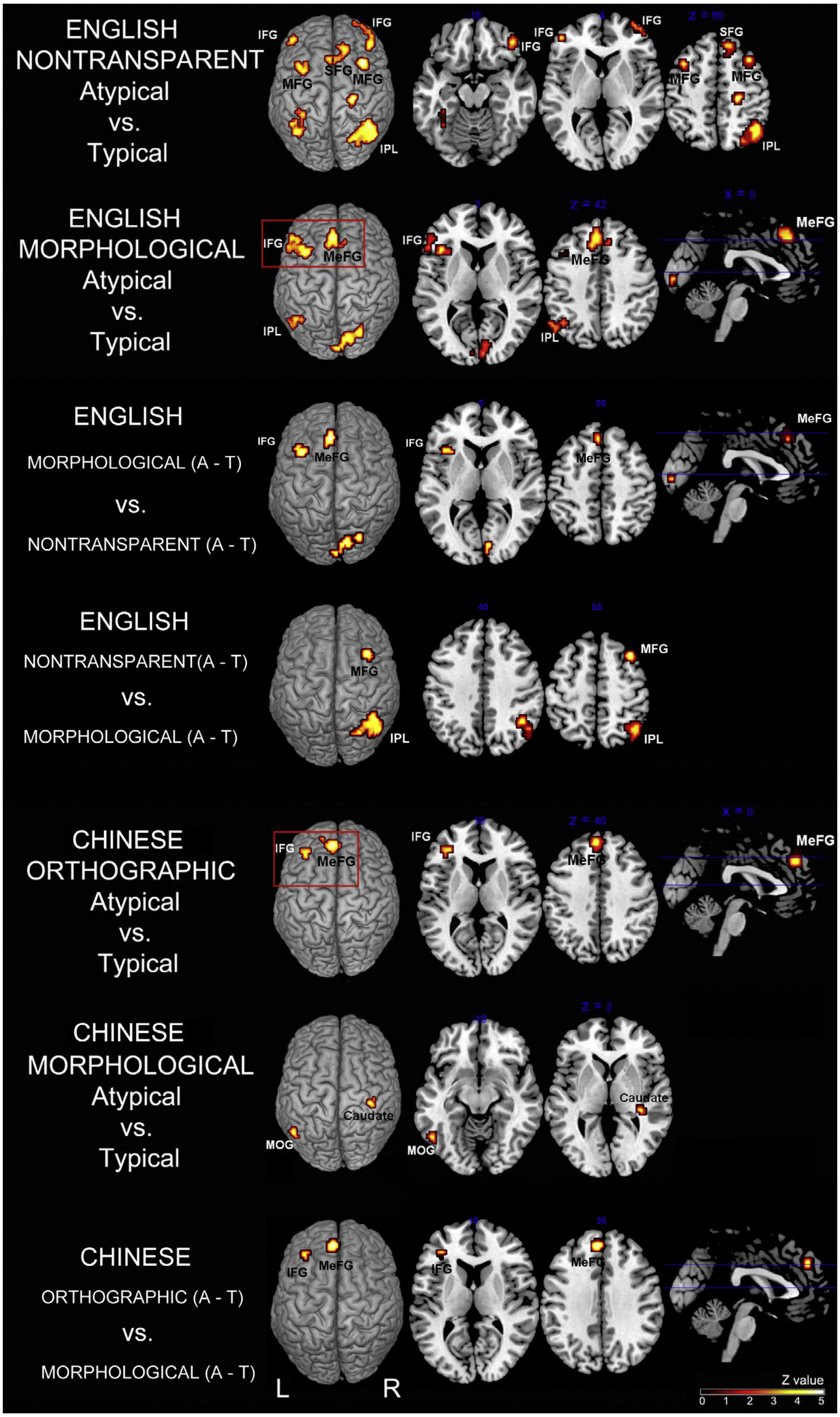

Fig. 4. English and Chinese speakers showed different patterns of activation for the typicality effect dependent on the type of linguistic label (Table 3). English nontransparent items and Chinese morphologically transparent items showed dissimilar activation in the bilateral frontal regions, even though 4 out of 5 categories had identical item pictures (Fig. 1). In contrast, English morphologically transparent and Chinese orthographically transparent items showed similar activation in the left IFG (BA46 and 47) and the left MeFG (BA8) (Fig. 5), even though all 5 categories contained item pictures completely different across the two languages (Fig. 1).

in frontal regions, whereas Chinese "orthographically transparent" items (e.g., BUG *chong2*虫: butterfly *hu2die2*蝴蝶) activated the left MeFG (BA8) and IFG (BA46, 47), two areas that overlap with activation for English morphologically transparent items (Fig. 4 and 5; Table 3).

In summary, despite overall similarities in the behavioral results and patterns of brain activation for the Yes *vs.* No contrast, English and Chinese speakers showed significant differences in brain activation patterns for the typicality effect. In English speakers, atypical items elicited larger activity in the left IFG, the right

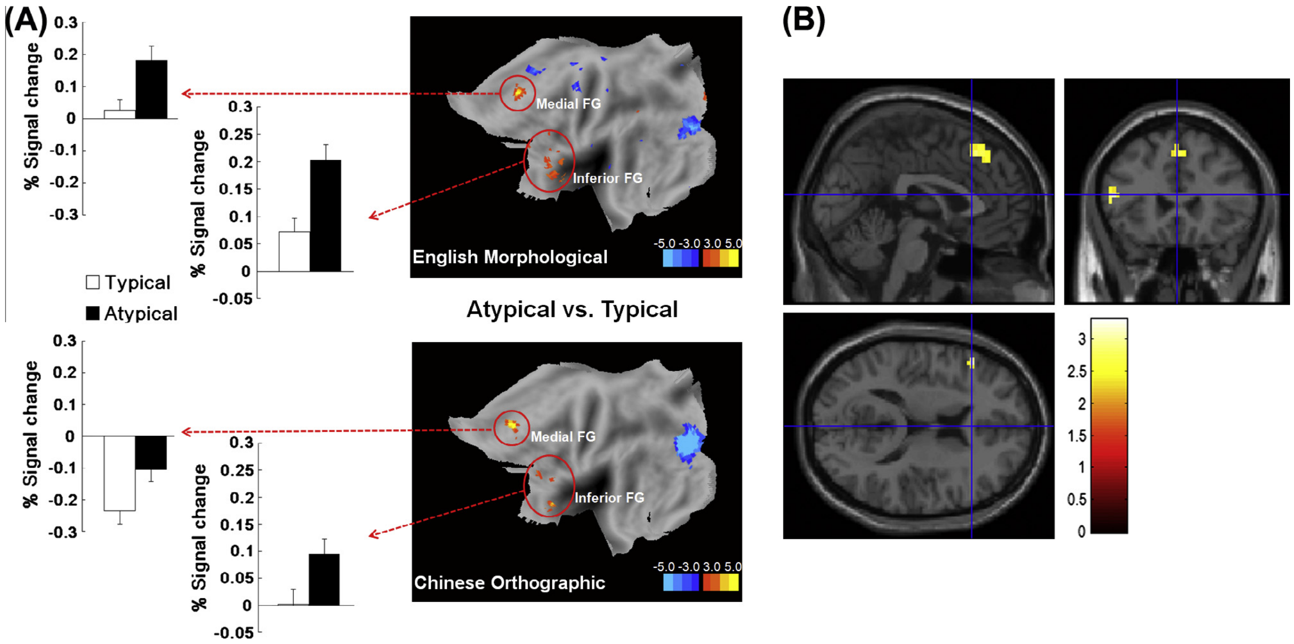

Fig. 5. Activation map and% Signal change for similar activation of the atypical vs. typical contrast in the left Inferior FG and Medial FG clusters for English morphologically transparent items and Chinese Orthographically transparent items (A). Conjunction analysis for the English morphologically transparent items and Chinese Orthographically transparent items revealed significant activation at the left Inferior FG and Medial FG clusters (B).

Table 3
Brain regions showing significant activation between the typical and atypical items for different label types in English and Chinese. (One sample *t*-test, uncorrected voxelwise threshold of $P < 0.005$, $P < 0.05$ FDR corrected for multiple comparisons with the small volume correction (SVC) in the regions of frontal and parietal cortex ($K > 20$ voxels). The threshold for the Chinese morphologically transparent condition (Atypical vs. typical) was set to $P < 0.005$, $P < 0.1$ with FDR corrected for multiple comparisons with the SVC in the left BA19 and right Caudate tail ($K > 10$ voxels)).

Contrast	BA	*P(FDR)*	Voxel	*x*	*y*	*z*	*Z*
English							
Morphological (Atypical–Typical)							
L Medial Frontal G	8	0.03	64	−6	35	38	3.80
L Inferior Frontal G	47	0.02	121	−38	18	−5	3.63
L Middle Frontal G	46	0.04		−41	43	7	2.72
Nontransparent (Atypical–Typical)							
L Superior Frontal G	8	0.02	31	−38	15	48	3.37
L Middle Frontal G	6	0.04		−28	21	52	2.75
L Inferior Frontal G	46	0.04	27	−47	43	7	3.43
R Superior Frontal G	8	0.02	39	16	35	47	3.42
L Superior Frontal G	8	0.03		−3	24	52	3.17
R Superior Frontal G	8	0.01	37	34	21	52	4.09
R Inferior Parietal L	40	0.02	182	44	−49	42	3.9
Morphological (Atypical–Typical)–Nontransparent (Atypical–Typical)							
Nontransparent (Atypical > Typical)							
L Medial Frontal G	8	0.03	24	−6	35	40	3.94
L Inferior Frontal G	47	0.02	28	−38	18	−5	3.63
Nontransparent (Atypical–Typical)–Morphological (Atypical–Typical)							
R Superior Frontal G	8	0.01	21	34	21	52	4.23
R Inferior Parietal L	40	0.01	105	47	−47	40	4.12
Chinese							
Morphological Atypical–Typical							
Atypical							
L Middle Occipital G	19	0.06	13	−49	−61	−5	3.51
R Caudate		0.06	12	38	−30	−3	2.64
Orthographic Atypical–Typical							
L Medial Frontal G	8	0.03	54	−3	47	38	3.92
L Inferior Frontal G	47	0.04	21	−28	18	−13	3.61
L Inferior Frontal G	46	0.04		−34	34	11	3.36
Orthographic (Atypical–Typical)–Morphological (Atypical–Typical)							
L Medial Frontal G	8	0.02	27	−3	47	38	3.98
L Inferior Frontal G	47	0.05	12	−26	18	−10	3.53
Morphological (Atypical–Typical)–Orthographic (Atypical–Typical)							
(none)							

MFG, the right SFG and the right IPL than typical items. None of these patterns of increased activation for atypical items were found in Chinese speakers. Further comparisons within each language demonstrate that brain activation for less typical items is inversely correlated with the prevalence and availability of linguistic cues to category membership, even when the items to be judged are pictorial stimuli. English items with nontransparent labels showed the most activation in the bilateral SFG, IFG, MFG and the right IPL, whereas English items with morphologically transparent labels and Chinese items with orthographically transparent labels showed the same activation in the left MeFG and IFG. Chinese items with morphologically transparent labels showed the least activation – only in the left MOG and the right caudate.

4. Discussion

In the current study, we investigated the representation of category typicality in the brain and its cross-linguistic variations with native US English and native Mandarin Chinese speakers. We conducted a category verification task with pictorial stimuli that differ in both typicality and linguistic cues to category membership. The results demonstrate that the availability of linguistic cues, even for pictorially presented stimuli, differentially affect brain activation during categorization of typical and atypical category members.

In the present study, the category verification task required that the participant first read and keep in mind a category label (e.g., vehicle) and then judge whether a picture (e.g., a sedan), shown 1500 ms later, was an example of the label. Because participants ultimately had to make a link between the visual characteristics of the picture and the linguistic category label shown before it, we assume that all participants, both Chinese and English speakers, engaged in some sort of semantic access before the final decision was made. However, whether the typicality of the item affected the semantic access or not, was strongly dependent on the availability of a linguistic cue to the category label, even when that linguistic cue was not explicitly presented.

As far as we know, no previous fMRI study has explored the representation of the "typicality effect" involved in semantic categorization. Nonetheless, our results are consistent with other studies involving semantic categorization. Specifically, categorizing both word and pictorial stimuli will elicit activation in the bilateral middle and inferior frontal gyrus, the bilateral inferior parietal lobule, bilateral temporal-occipital conjunctions, anterior cingulate cortex and caudate (Adams & Janata, 2002; Ganis et al., 2007; Grossman et al., 2002; Jiang et al., 2007; Koenig et al., 2005; Myers, 2007; Reber et al., 2003). One particularly important region is the left inferior frontal gyrus (IFG), which has been found to contributes greatly to semantic and lexical access for English speakers (Bookheimer, 2002; Hagoort et al., 2004; Vandenberghe et al., 1996) as well as bilinguals (Rodriguez-Fornells, Rotte, Heinze, Nosselt, & Munte, 2002) and monolingual Chinese speakers (Liu et al., 2008; Siok, Perfetti, Jin, & Tan, 2004). Our study echoes these findings of greater left IFG activation for both the English- and Chinese speaking participants for out-of-category items (*No* responses) relative to in-category items (*Yes* responses). Moreover, in English, atypical items showed larger left IFG activity than typical items, regardless of whether the verbal label for that item was nontransparent or whether it contained a morphological cue to its category membership. In Chinese, atypical items showed larger left IFG activation than typical items, but only for items with orthographic cues to category membership (Fig. 4B, 5; Table 3). These results greatly underscore the role of the left IFG in semantic processing and demonstrate it might be a fundamental locus of the typicality effect in semantic categorization.

However, there is also another possible explanation on the role of the left IFG in categorization. According to some language comprehension models, such as MUC model (Hagoort, 2005), the left IFG is the areas of semantic unification whereas the temporal lobe is the areas underlying lexical and semantic retrieval (Willems, Ozyurek, & Hagoort, 2007; Zhu et al., 2012). Moreover, some studies on visual processing also demonstrated that when there was conflict between two sources of information, the left IFG was activated, as compared with no conflict. It seems that the more difficult the unification processing, the more possible that the left IFG was activated. Thus in the present study it might be that the left IFG activity for atypical item was mainly due to the fact that, relative to typical item, atypical item was more difficult to match with the corresponding "category" (as reflected by the behavioral results). Meanwhile, English speakers responded more slowly for pictures with morphologically transparent items than nontransparent items, whereas Chinese speakers responded more slowly for Orthographic than Morphological items. Therefore, the activation of left IFG was more pronounced in the morphologically transparent item for English speakers whereas more pronounced in the orthographically transparent items for Chinese speakers. Further studies thus are needed to verify these two possible roles of the left IFG in categorization processing.

In addition to the left IFG, both groups of participants also showed a typicality effect in the bilateral MeFG, areas which have been argued to be related to goal-directed attention (Corbetta & Shulman, 2002), decision making, and category uncertainty (Grinband, Hirsch, & Ferrera, 2006). Activation in these areas is consistent with behavioral results, showing typicality effects in both accuracy and reaction time for Chinese as well as English speakers. More specifically, they also suggest that the "typicality effect", at least for semantic category judgments, may reside largely in frontal areas responsible for retrieval from semantic memory and working memory (e.g., Badre, Poldrack, Pare-Blagoev, Insler, & Wagner, 2005; Feredoes & Postle, 2007) rather than the parietal areas that were also active during this particular task which involved integrating visual with semantic information.

Our results thus showed that the left IFG and the bilateral MeFG are the most consistent neural correlates of typicality effect that persist across languages, which could reflect the semantic processing and explicit decision making processing that accompany categorization.

The second question of interest is whether participants in the two languages engaged in the same type of semantic access and/or whether they engaged in additional semantic processing pictorial stimuli that were atypical members of a category, given the differences in the availability of linguistic cues to category membership (e.g., car/*jiao4che1*轿车). Our assumption, based on the present data, is that speakers of English not only accessed a verbal label for the pictures, but that they engaged in additional semantic and phonological processing, evidenced by the presence of widespread activation for the Nontransparent Atypical items in the IFG (bilaterally) and MFG (primarily left hemisphere) (Fig. 4A), in order to facilitate judgments in this task. In contrast, speakers of Chinese were able to bypass much of this additional semantic processing because of the presence of linguistic cues in the common label for morphologically transparent items (e.g., *jiao4che1*轿车) and the prevalence of this naming convention in Chinese. Thus, because they were able to more directly access category information through the morphological information available in the common labels for these pictures, they did not show a typicality effect in frontal regions, although they still engaged in visual categorization processes in the middle occipital gyrus (Pernet et al., 2004) and caudate (Grossman et al., 2002) (Fig. 4D), as would be expected for implicit rule-based categorization (Grossman et al., 2002).

Interestingly, however, when English participants were given pictures with labels that also contained a morphological cue to the category, the same type of semantic processing bypass did not occur, and we argue that this is because this type of cue is not a consistent marker of category membership in English. Nonetheless, it did appear to help with semantic and phonological processing, as evidenced by reduced activation in the right IFG for the morphologically transparent condition relative to the nontransparent condition (Fig. 4B). Nonetheless, providing a morphological cue to category membership in English might not be as effective as it is in Chinese, a language which has had a long tradition of using linguistic cues to category for thousands of years across large numbers of semantic categories.

Moreover, it is also possible that because the semantic radicals in Chinese orthographically transparent items are not pronounced and thus do not provide additional phonological information in the way that morphologically transparent items do, it may be that they simply provide less explicit and later-learned cues than morphologically transparent items. As a result, Chinese speakers may have to evoke additional semantic processing in the left IFG when viewing orthographically transparent items (Fig. 4C). This finding is consistent with ERP studies using this same paradigm, in which Chinese speakers did not show a reliable typicality effect in the N300 and N400 components for items with morphological cues to category membership, whereas they did show significant typicality effects in both the N300 and N400 components for items with orthographic cues to category membership (Liu et al., 2010). A recent finding in a semantic relatedness judgment task, where the left ventral IFG (BA47), the same area that was activated in the orthographic Atypical-Typical comparison in our category judgment task, also shows such activation to be related to semantic, but not phonological, processing in Chinese speakers (Liu et al., 2008).

The reduction of frontal activity for picture stimuli with morphologically transparent labels, relative to those with nontransparent labels for English speakers, and the increase in frontal activation for stimuli with orthographically transparent labels in Chinese speakers, together demonstrate the influence of multiple levels of linguistic cues in categorization, and the importance of the prevalence and explicitness of such cues in the language. These results are in line with previous findings about different categorization systems in the brain (Smith & Grossman, 2008) as well as with the proposal that these processes are widely distributed across brain areas (Rogers & McClelland, 2004). For English nontransparent items with no explicit linguistic cues, speakers rely on semantic rule-based categorization with a loading on working memory and selective attention, resulting in bilateral inferior frontal gyrus and medial gyrus activation. In contrast, for morphologically transparent items in English and orthographically transparent items in Chinese, for which some category-relevant information is available – but not highly prevalent or explicit – speakers need to conduct on-line linguistic rule-based categorizations (e.g., the morphological cues in English and orthographical cues in Chinese) requiring less effort, and resulting in left-lateralized inferior frontal gyrus activation. Nonetheless, with follow-up studies focused more specifically on delineating the differences between these types of cues, more precise distinctions might be found between the English morphologically transparent and Chinese orthographically transparent items.

Our findings have enriched our understanding of categorization and semantic processing and will allow us to develop more parsimonious theories to explain these fundamental human abilities from neurological and cross-cultural perspectives. Although traditional behavioral theories such as the "probabilistic" and "exemplar" views (Medin & Smith, 1984) do not provide specific predictions on how the linguistic cues in a language could influence categorization and semantic processing in the brain, probably because of the lack of sensitivity of those behavioral methods, the PDP approach (Rogers & McClelland, 2004, 2008) does make explicit predictions on how category information in semantic memory is distributed in the brain. According to Rogers and McClelland (2004), the content of semantic memory is represented in the same regions of cortex that directly encode modality-specific regularities in the environment during perception and action. However, domain-general learning mechanisms operate to allow the semantic system, when presented with information about an object in some perceptual modality (e.g., visual or auditory), to make correct inferences about the object's unspecified attributes. As a consequence, the system acquires abstract representations whose similarity relations are not tied to any individual modality (e.g., visual or auditory), but capture the deep structure across modalities, most likely in the frontal regions. Most importantly, the maturation of this PDP representation is highly dependent on the training process in the neural network. In computer modeling, training is achieved by running more epochs, whereas in reality, training is achieved by accumulating more experiences with the items and their properties such as linguistic cues. Our current fMRI data support the critical role of training and experience on forming representations in semantic memory by showing a pattern such that the morphologically transparent items for English speakers and orthographically transparent items for Chinese speakers, when requiring similar experiences and knowledge to master their categorical cues in corresponding language, actually activate similar brain areas in the inferior and medial frontal gyrus.

One limitation of the current study is that we did not include non-transparent Chinese items because of the difficulty in finding proper stimuli that could be matched with the typicality and visual complexity characteristics both between and across languages. All the stimuli used in the study were adopted from our previous one with two pilot studies of category replacing and picture naming (Liu et al., 2010). However, non-transparent items are much less common than linguistically labeled items in Chinese and those we found failed badly in both of these two pilot studies. As a result, we only included two "non-transparent" categories in our previous cross-cultural ERP studies (STATION and BUILDING), but because they were both much more visually complicated than other stimuli and showed poor accuracy in the behavioral results, we discarded them in the present study. For the above reasons, we could not find enough Chinese "non-transparent" items and thus could not examine this interesting condition in the current study.

Of interest for future research, therefore, is whether and what types of differences might appear due to the specific type of information provided (orthographic vs. morphological) vs. the explicitness of the information, or simply the pervasiveness of the information in the language (morphological transparency is a dominant feature of most Chinese nouns but a less common, though still present, feature of some English nouns). Nonetheless, it is striking with all these differences across languages that the patterns were so similar between morphologically cued items in English and orthographically cued items in Chinese. To answer these questions, a promising line of research will be including a truly bilingual Chinese-English group (e.g., participants from Hong Kong or Singapore) and testing the influence of language on their categorization.

Finally, for Chinese morphologically transparent items, automated and direct access to semantic and phonological components in implicit long-term memory appear to be possible, for which explicit rule-based categorization processes do not appear to be necessary. This possibility raises a number of questions about the role of typicality and categorization processes more generally. Specifically, despite similar increases in reaction time for atypical (e.g., "ostrich") relative to typical (e.g., "robin") members of a category in both English and Chinese speakers, the brain does not necessarily process typicality in similar ways across languages, at least when it comes to deciding on category membership. These data suggest, further, that typicality is a useful heuristic only when a language does not regularly embed category-level terms in the labels for members of the category.

Most importantly, however, these data speak to larger issues of how similar behavioral results can obtain despite quite dissimilar underlying brain processes. Both the similarities and the differences between English and Chinese speakers on this categorization task speak to the flexibility and complexity of brain processes underlying apparently similar behavioral responses. Our data suggest that these differences that may impact a number of processes in which typicality plays a role (e.g., in the behavioral and brain manifestations also of semantic dementia), but they may also be just one of many phenomena in which the neurophysiological underpinnings of common cognitive processes may inform important differences in how language, and experience more generally, shapes the brain.

Acknowledgments

We thank Yuanyuan Gu for helping collecting the fMRI data. We also thank Dr. Daniel Weissman for comments and suggestions at various stages of this project. This work was supported by the National Basic Research Program of China (2011CB711000, 2013CB837300), the National Natural Science Foundation of China (NSFC) (31170971, 61210010), Major Project of National Social Science Foundation (12&ZD228) to C.L. and CNLYB1212 to X.M., as well as the National Science Foundation grant BCS-0350272 and internal funding from the Center for Human Growth and Development, National Institutes of Health training grant TW000035-07S1 (Minority International Research Training) to T.T.

References

Adams, R. B., & Janata, P. (2002). A comparison of neural circuits underlying auditory and visual object categorization. *Neuroimage, 16*(2), 361–377.

Ameel, E., Malt, B. C., Storms, G., & Van Assche, F. (2009). Semantic convergence in the bilingual lexicon. *Journal of Memory and Language, 60*(2), 270–290.

Ameel, E., Storms, G., Malt, B. C., & Sloman, S. A. (2005). How bilinguals solve the naming problem. *Journal of Memory and Language, 53*(1), 60–80.

Badre, D., Poldrack, R. A., Pare-Blagoev, E. J., Insler, R. Z., & Wagner, A. D. (2005). Dissociable controlled retrieval and generalized selection mechanisms in ventrolateral prefrontal cortex. *Neuron, 47*(6), 907–918.

Barrett, S. E., & Rugg, M. D. (1990). Event-related potentials and the semantic matching of pictures. *Brain and Cognition, 14*(2), 201–212.

Bookheimer, S. (2002). Functional MRI of language: New approaches to understanding the cortical organization of semantic processing. *Annual Review of Neuroscience, 25*, 151–188.

Brett, M., Anton, J. -L., Valabregue, R., & Poline, J. -B. (2002). Region of interest analysis using an SPM toolbox. In *Paper presented at the 8th international conference on functional mapping of the human brain*, 2002 June 2–6; Sendai, Japan. Neuroimage 16: CD-ROM.

Canli, T., Omura, K., Haas, B. W., Fallgatter, A., Constable, R. T., & Lesch, K. P. (2005). Beyond affect: A role for genetic variation of the serotonin transporter in neural activation during a cognitive attention task. *Proceedings of the National Academy of Sciences of the United States of America, 102*(34), 12224–12229.

Corbetta, M., & Shulman, G. L. (2002). Control of goal-directed and stimulus-driven attention in the brain. *Nature Reviews Neuroscience, 3*(3), 201–215.

Depue, B. E., Curran, T., & Banich, M. T. (2007). Prefrontal regions orchestrate suppression of emotional memories via a two-phase process. *Science, 317*(5835), 215–219.

Feredoes, E., & Postle, B. R. (2007). Localization of load sensitivity of working memory storage: Quantitatively and qualitatively discrepant results yielded by single-subject and group-averaged approaches to fMRI group analysis. *Neuroimage, 35*(2), 881–903.

Friston, K. J., Ashburner, J., Frith, C. D., Poline, J. B., Heather, J. D., & Frackowiak, R. S. J. (1995). Spatial registration and normalization of images. *Human Brain Mapping, 3*(3), 165–189.

Fujihara, N., Nageishi, Y., Koyama, S., & Nakajima, Y. (1998). Electrophysiological evidence for the typicality effect of human cognitive categorization. *International Journal of Psychophysiology, 29*(1), 65–75.

Ganis, G., Schendan, H. E., & Kosslyn, S. M. (2007). Neuroimaging evidence for object model verification theory: Role of prefrontal control in visual object categorization. *Neuroimage, 34*(1), 384–398.

Grinband, J., Hirsch, J., & Ferrera, V. P. (2006). A neural representation of categorization uncertainty in the human brain. *Neuron, 49*(5), 757–763.

Grossman, M., Smith, E. E., Koenig, P., Glosser, G., DeVita, C., Moore, P., et al. (2002). The neural basis for categorization in semantic memory. *Neuroimage, 17*(3), 1549–1561.

Hagoort, P. (2005). On Broca, brain, and binding: A new framework. *Trends in Cognitive Sciences, 9*(9), 416–423.

Hagoort, P., Hald, L., Bastiaansen, M., & Petersson, K. M. (2004). Integration of word meaning and world knowledge in language comprehension. *Science, 304*(5669), 438–441.

Hauk, O., Patterson, K., Woollams, A., Cooper-Pye, E., Pulvermuller, F., & Rogers, T. T. (2007). How the camel lost its hump: The impact of object typicality on event-related potential signals in object decision. *Journal of Cognitive Neuroscience, 19*(8), 1338–1353.

Heinze, H. J., Muente, T. F., & Kutas, M. (1998). Context effects in a category verification task as assessed by event-related brain potential (ERP) measures. *Biological Psychology, 47*(2), 121–135.

Jiang, X., Bradley, E., Rini, R. A., Zeffiro, T., VanMeter, J., & Riesenhuber, M. (2007). Categorization training results in shape- and category-selective human neural plasticity. *Neuron, 53*(6), 891–903.

Koenig, P., Smith, E. E., Glosser, G., DeVita, C., Moore, P., McMillan, C., et al. (2005). The neural basis for novel semantic categorization. *Neuroimage, 24*(2), 369–383.

Lancaster, J. L., Rainey, L. H., Summerlin, J. L., Freitas, C. S., Fox, P. T., Evans, A. C., et al. (1997). Automated labeling of the human brain: A preliminary report on the development and evaluation of a forward-transform method. *Human Brain Mapping, 5*(4), 238–242.

Liu, C., Tardif, T., Mai, X. Q., Gehring, W. J., Simms, N., & Luo, Y. J. (2010). What's in a name? Brain activity reveals categorization processes differ across languages. *Human Brain Mapping, 31*(11), 1786–1801.

Liu, C., Zhang, W. T., Tang, Y. Y., Mai, X. Q., Chen, H. C., Tardif, T., et al. (2008). The visual word form area: Evidence from an fMRI study of implicit processing of Chinese characters. *Neuroimage, 40*(3), 1350–1361.

Maldjian, J. A., Laurienti, P. J., Kraft, R. A., & Burdette, J. H. (2003). An automated method for neuroanatomic and cytoarchitectonic atlas-based interrogation of fMRI data sets. *Neuroimage, 19*(3), 1233–1239.

Martinez, I., & Shatz, M. (1996). Linguistic influences on categorization in preschool children: A cross linguistic study. *Journal of Child Language, 23*, 529–545.

McPherson, W. B., & Holcomb, P. J. (1999). An electrophysiological investigation of semantic priming with pictures of real objects. *Psychophysiology, 36*(1), 53–65.

Medin, D. L., & Smith, E. E. (1984). Concepts and concept-formation. *Annual Review of Psychology, 35*, 113–138.

Mervis, C. B., Catlin, J., & Rosch, E. (1976). Relationships among goodness-of-example, category norms, and word-frequency. *Bulletin of the Psychonomic Society, 7*(3), 283–284.

Mervis, C. B., & Rosch, E. (1981). Categorization of natural objects. *Annual Review of Psychology, 32*, 89–115.

Myers, E. B. (2007). Dissociable effects of phonetic competition and category typicality in a phonetic categorization task: An fMRI investigation. *Neuropsychologia, 45*(7), 1463–1473.

Patterson, K. (2007). The reign of typicality in semantic memory. *Philosophical Transactions of the Royal Society B-Biological Sciences, 362*(1481), 813–821.

Pernet, C., Franceries, X., Basan, S., Cassol, E., Demonet, J. F., & Celsis, P. (2004). Anatomy and time course of discrimination and categorization processes in vision: An fMRI study. *Neuroimage, 22*(4), 1563–1577.

Reber, P. J., Gitelman, D. R., Parrish, T. B., & Mesulam, M. M. (2003). Dissociating explicit and implicit category knowledge with fMRI. *Journal of Cognitive Neuroscience, 15*(4), 574–583.

Rodriguez-Fornells, A., Rotte, M., Heinze, H. J., Nosselt, T., & Munte, T. F. (2002). Brain potential and functional MRI evidence for how to handle two languages with one brain. *Nature, 415*(6875), 1026–1029.

Rogers, T. T., & McClelland, J. L. (2008). Precis of semantic cognition: A parallel distributed processing approach. *Behavioral and Brain Sciences, 31*(6), 689–749.

Rogers, T. T., & McClelland, J. L. (2004). *Semantic cognition: A parallel distributed processing approach*. Cambridge, MA: MIT Press.

Seymour, B., Daw, N., Dayan, P., Singer, T., & Dolan, R. (2007). Differential encoding of losses and gains in the human striatum. *Journal of Neuroscience, 27*(18), 4826–4831.

Siok, W. T., Perfetti, C. A., Jin, Z., & Tan, L. H. (2004). Biological abnormality of impaired reading is constrained by culture. *Nature, 431*(7004), 71–76.

Smith, E. E., & Grossman, M. (2008). Multiple systems of category learning. *Neuroscience and Biobehavioral Reviews, 32*(2), 249–264.

Stuss, D. T., Picton, T. W., & Cerri, A. M. (1988). Electrophysiological manifestations of typicality judgment. *Brain and Language, 33*(2), 260–272.

Talairach, J., & Tournoux, P. (1988). Co-planar stereotaxic atlas of the human brain. 3-Dimensional proportional system: an approach to cerebral imaging.

Tardif, T. (2006). *But are they really verbs?: Chinese words for action.* Oxford.

Vandenberghe, R., Price, C., Wise, R., Josephs, O., & Frackowiak, R. S. J. (1996). Functional anatomy of a common semantic system for words and pictures. *Nature, 383*(6597), 254–256.

Whorf, B. L. (1956). *Language, thought and reality: Selected writings of Benjamin Lee Whorf.* Cambridge, MA: MIT Press.

Willems, R. M., Ozyurek, A., & Hagoort, P. (2007). When language meets action: The neural integration of gesture and speech. *Cerebral Cortex, 17*(10), 2322–2333.

Woollams, A. M., Silani, G., Okada, K., Patterson, K., & Price, C. J. (2011). Word or word-like? Dissociating orthographic typicality from lexicality in the left occipito-temporal cortex. *Journal of Cognitive Neuroscience, 23*(4), 992–1002.

Yoshida, H., & Smith, L. B. (2003). Known and novel noun extensions: Attention at two levels of abstraction. *Child Development, 76*(2), 564–577.

Zhou, Y. G. (1978). To what degree are the "phonetics" of present-day Chinese characters still phonetic? *Zhongguo Yuwen, 146*, 172–177.

Zhou, X. L., Marslen-Wilson, W., Taft, M., & Shu, H. (1999). Morphology, orthography, and phonology in reading Chinese compound words. *Language and Cognitive Processes, 14*(5–6), 525–565.

Zhu, Z. D., Hagoort, P., Zhang, J. X., Feng, G. Y., Chen, H. C., Bastiaansen, M., et al. (2012). The anterior left inferior frontal gyrus contributes to semantic unification. *Neuroimage, 60*(4), 2230–2237.