Chapter I

Introduction

Human adults frequently treat apparently different properties, objects, or events as effectively similar; that is, they categorize. Categorization involves the grouping of separate items into a set according to some type of precept or rule (Bornstein, 1984; Gelman & Diesendruck, 1999; Mandler & McDonough, 1998). The internal structure of category knowledge in semantic memory is assumed to have organizing principles in human adults (Rosch, 1973), among which an important one is the typicality effect. *Is a canola a flower? Is a rose a flower?* When asked questions such as this, North American English speakers generally respond to "rose" more quickly and more accurately than "canola" (*Brassica napus*, or the flower used in canola oil). Many researchers (Bjorklund, Thompson, & Ornstein, 1983; Boster, 1988; Chumbley, 1986; Komatsu, 1992; Mervis, Catlin, & Rosch, 1976; Mervis & Rosch, 1981; Schwanenflugel & Rey, 1986) believe that this is because most people think a rose is a more typical example of the category "flower" than canola. Such differences in typicality have been found to affect categorization processes (Rosch et al., 1976). Behavioural studies show that typical members appear to be privileged over atypical members. For example, reaction time is faster in category judgement tasks for typical exemplars than atypical exemplars (for a review see Danks & Glucksberg, 1980), and more typical exemplars are named than atypical exemplars when people are asked to spontaneously produce lists of category members (Mervis, Catlin, & Rosch., 1976). Moreover, both children (Mervis & Pani, 1980) and adults (Rosch, 1973a, b) learn more representative exemplars of a category before learning less representative exemplars.

As one of the most consistent indexes of categorization in semantic memory in behavioral studies (Mervis & Rosch, 1981), the typicality effect has also been investigated with neurophysiological techniques such as the Event-Related Potential (ERP), the scalp-recorded changes in electrical activity that occur in response to a sensory, cognitive or motor event. In contrast to behavioral measures, ERP measures provide a millisecond-by-millisecond record of the electrical activity that occurs in the brain during the process of interest (Osterhout et al., 1997). Moreover, the typicality effect also has the potential to be investigated with neuroimaging techniques with high spatial resolution such as functional Magnet Reasoning Imaging (fMRI) (a technique used to visualize brain function in 3D dimension by visualizing changes in chemical composition of brain areas or changes in the flow of fluids that occur over timespans of seconds to minutes). Finally, although the typicality effect has been found in behavioural measures with children (Anglin, 1977; Mervis & Pani, 1980), how exactly it is shaped and developed in children is still unclear, and few studies have began investigate the neurophysiclogical developmental of categorization. This would allow us to not only understand the time course of categorization, but to also understand which parts of the brain are involved in these processes and how we might address issues in patients with semantic memory and other deficits relevant on categorization.

The ERP component that has been repeatedly identified in the ERP study of typicality effects is the N400, a component well known to be an index of syntax and semantic processing in language studies (Kutas & Hillyard, 1980, Kutas, & Hillyard, 1984). In a study that investigated typicality judgments of words that were of either high or low frequency, Stuss and Cerri (1988) found an N400 component correlated with the goodness-of-fit of a word to a particular category independent of frequency, with poor examples of the category evoking a significant more negative N400 than good examples. Fujihara and his colleagues (Fujihara, Nageishi, Koyama, & Nakajima, 1998) conducted an ERP study to investigate the typicality effect in Japanese. In a category verification task with stimulus words that were either typical (e.g. "carrot"), or atypical (e.g. "parsley") members of a given category (e.g. "vegetables"), participants were required to judge whether or not each stimulus belonged to a target category (e.g., "vegetables" or "sports"). The results showed that, for the target category, typical words were responded to more quickly than were atypical words and the ERP amplitudes between a 300-450 ms period (N300) were more negative after atypical words than after typical words (i.e., they generated a “typicality effect”), which suggested that typical words (e.g. "carrot") were more primed by the target category label (e.g. "vegetables") than were atypical words (e.g. "parsley"). Another study with German- and English-speaking participants reported a similar result (Heinze, Muente, &Kutas, 1998). In their study, pairs of words (e.g., fish–whale, bird–bat) were presented and subjects were asked to decide whether or not the second word was an exemplar of the superordinate category indicated by the first word. Both German and English subjects yielded similar results. Moreover, false exemplars (i.e., out of category items) elicited the largest N400s. In addition, ERPs for the true exemplars were influenced by typicality. Atypical items yielded larger N400s than typical items. Such a typicality effect was invariant in the face of a context manipulation (i.e., regardless of whether the false item was *related* to the true item, e.g. bird-bat, or *unrelated*, e.g., bird-car). These results suggest that category membership decision and the typicality of an exemplar (as reflected in N400 component) can vary independently from each other and, therefore, likely index the activity of different brain systems. In addition to the N400 component found in semantic processing of word stimuli, studies with pictorial stimuli have found an additional positive component 160 milliseconds (P160) and a negative component 300 milliseconds (N300) post -stimulus that represent the perceptual and semantic processing of pictorial objects (Barrett & Rugg, 1990; Hauk et al., 2007; Kiefer, 2001; McPherson & Holcomb, 1999). For example, semantically unrelated object pairs (e.g., chair and cat) elicit larger N300 and N400 components than semantically related object pairs (e.g., hamburger and chips) in a relatedness judgment task (McPherson & Holcomb, 1999).

Although the classic “typicality effect” has yet not been directly investigated using fMRI, categorization studies using fMRI with English normal participants and patients have identified three qualitatively different categorization systems in the brain (Ashby & Casale, 2005; Grossman et al., 2003; Grossman et al., 2006; Grossman et al., 2002; Koenig et al., 2002; Koenig et al., 2005; Koenig, Smith, Moore, Glosser, & Grossman, 2007; Koenig et al., 2008; Patalano, Smith, Jonides, & Koeppe, 2001). First, a rule-based categorization process has been associated with the working memory system and selective attention in the frontal and parietal areas, especially the left inferior frontal gyrus. Second, a similarity-based categorization process has been associated with explicit long term memory and integration of perceptual features in the parietal-temporal areas. Third, other sorts of implicit categorization processes have been associated with implicit long term memory in the temporal-occipital areas (for review, see (E. E. Smith & Grossman, 2008)). However, these hypotheses about different categorization systems are mainly obtained from fMRI data. We don't know how these system differ in the time course of categorization processing. Thus an integrated model with both spatial and temporal descriptions is necessary to fully demonstrate the categorization processing in the brain.

Different theories of semantic memory and categorization have different interpretations of the typicality effect. There are three developed views about how concepts are categorized in human mind: "classical," "probabilistic," and "exemplar" views (Medin & Smith, 1984). The *classical view* holds that all instances of a concept share common properties that are necessary and sufficient conditions for defining the concept. The *probabilistic view* denies that there are defining properties, and instead argues that concepts are represented in terms of properties that are only characteristic or probable of class members. Membership in a category can thus be graded rather than all-or-none, where the better members have more characteristic properties than the poorer ones. The *exemplar view* agrees with the claim that concepts need not contain defining properties, but further claims that categories may be represented by their individual exemplars, and that assignment of a new instance to a category is determined by whether the instance is sufficiently similar to one or more of the category's known exemplars. We will focus only on the probabilistic and exemplar view because the classical view cannot explain the typicality effect since the essence of the typicality effect demonstrates that not all members of a category have equal status. Instead, exemplars judged to be typical of a concept are generally categorized faster and more accurately than exemplars judged less typical (Rosch & Mervis, 1975).

The probabilistic view assumes that categories are abstractions, or summary representations, but argues that for a property to be included in the summary it need have only a substantial probability of occurring in instances of the category, i.e. it need only be characteristic of the category, not defining (Mervis et al., 1976; Rosch, 1973). An object will then be categorized as an instance of some category A if, for example, it possesses some criterial number of properties, or sum of weighted properties, included in the summary representation of A. Categorization is thus a matter of assessing similarity rather than of applying a definition. The probabilistic view is thus able to address typicality effects. Items are typical of a category to the extent that they contain properties that are characteristic of the category. This idea makes typicality a variation of similarity. The more similar an item to a category, the faster and more reliably it can be judged to exceed some threshold level of similarity, resulting in the effects of typicality on categorization (Smith & Medin, 1981).

In contrast, the exemplar view assumes that, at least in part, any given category consists of separate descriptions of its exemplars. Some exemplar models allow for a more abstract representation as well (Medin & Schaffer 1978), but others (e.g. the average distance model evaluated by Reed, 1972) are based only on exemplars. Exemplar models have in common the idea that categorization of an object relies on comparisons of that object to known exemplars of the category. According to the exemplar view, typicality effects may arise because people are more likely to represent only typical members (Mervis 1980), or because typical instances are more similar to other stored exemplars. Being more similar to other stored exemplars of a category, typical instances should facilitate retrieval of exemplars from that category and hence be categorized more quickly and accurately.

Both the probabilistic and the exemplar view emphasize the role of similarity in shaping the typicality effect. However, neither theory clearly explains the relationship between similarity and other psycholinguistic factors such as familiarity and word frequency (E.E. Smith & Medin, 1981). In addition, the similarity account of the typicality effect needs to explain cross-linguistic and cross-cultural differences and similarities found in numerous empirical studies (Diesendruck & haLevi, 2006; Lin & Schwanenflugel, 1995; Russell, 1991; Schunn & Vera, 2004). If we aim for maximal parsimony in our theories of categorization, multiple mechanisms that vary across languages usually imply less parsimony. A parsimonious theory that allows for a simple mechanism to account for multiple findings across different elements (language and culture) is thus more desirable to explain categorization. The current set of studies is aimed both to provide fundamental data that must be accounted for in such a theory, and to help us better form such a theory.

In addition to these two cognitive theories originating from behavioural studies, a computational approach to categorization and semantic memory more generally is also worth considering. Based on the parallel distributed processing (PDP) approach (Rumelhart, McClelland, & Group, 1986) and the semantic memory models of Rumelhart (Rumelhart, 1990; Rumelhart & Todd, 1993) and Hinton (1981; 1986), Rogers and McClelland (2004) proposed that categorization is not the only efficient mechanism for storing and generalizing knowledge about the world. Instead, a connectionist approach, which they name "semantics without categorization" (Rogers and McClelland, 2010), requiring neither a new representational element for every new class or object, nor any internal process of categorization, is argued to account for the full variety of behavioural phenomena found in previous categorization studies, including the typicality effect. According to this approach, semantic representations do not need to extract, store, and retrieve attributes, facts, or propositions about objects; instead, they need only to allow such information to be produced as overt responses in particular task contexts (T. T. Rogers et al., 2004). In addition, it also provides some clues as to how the semantic system may be organized in the brain by suggesting that abstract semantic representations emerge as a product of statistical learning mechanisms in a region of the cortex suited to performing cross-modal mappings by virtue of its many interconnections with different perceptual-motor areas (McClelland, Rogers, Patterson, Dilkina, & Lambon Ralph, 2009; T. T. Rogers & McClelland, 2004, 2008; T. T. Rogers & McClelland, 2010). The key features of this approach are: 1) Cognitive phenomena arise from the propagation of activation amongst simple, neuron-like processing units; 2) Propagation of activation is constrained by weighted synapse-like connections between units; 3) Changes to weights in the system are generated by a process of predictive error-driven learning. (Rogers and McClelland, 2010). Using this PDP approach, McClelland and colleagues have successfully simulated impairments of the behavioural typicality effect in semantic dementia (SD) patients with a model in which semantic representations emerge from mechanisms that acquire the mappings between visual representations of objects and their verbal descriptions (T. T. Rogers et al., 2004). Theoretically, the PDP approach has great potential for much greater parsimony and for ways to explore both developmental and cross-cultural and linguistic differences, as well as how these processes are instantiated in the brain.

In a simple version of the PDP model of semantic memory (Farah and McClelland, 1991)(Fig. 1), there are three main layers of units, corresponding to verbal inputs (e.g., names), visual inputs (e.g., pictures), and semantic memory representations. The semantic memory units are divided into visual units and functional units. There are bidirectional connections between units both within and between layers, with the exception that there are no direct connections between the name and picture input units. Unlike these two traditional views of categorization and the hypothesis of multi-systems categorization systems in the brain, this PDP model provide explicit prediction about how and where the input auditory and visual information are integrated in the brain. However, it still did not explain the time course involved in the categorization processing. Again, to explain how semantic memory and categorization was process in the brain, we need a model integrated with both brain areas activated for the processing and the time course of these activation.

However, neither the PDP model nor the other theories address when typicality is in fact a useful cue for categorization and when or whether other types of cues might supersede it. In studies involving both experts and non-experts, for instance, only non-experts (US undergraduates vs. Itza' Mayans or US bird experts) relied on typicality to make judgments about bird classifications, which suggests that typicality is used as a "crutch" for categorization when other information is lacking (Bailenson, Shum, Atran, Medin, & Coley, 2002). In addition, the linguistic relativity hypothesis proposed by Whorf (1956) argues that the language one speaks influences the way one thinks. Would Chinese speakers then be as fast and accurate to classify canola *you2cai4hua1* 油菜花 as rose *mei2gui4hua1* 玫瑰花 because both words have the category term flower *hua1* 花 embedded in their names? In the present studies we explore the role of explicit linguistic labels as an alternative to a reliance on typicality during a categorization task.

Both Mandarin Chinese and English can provide explicit linguistic labels to category membership (e.g., canola *you2cai4hua1*油菜花 or rose *mei2gui4hua1*玫瑰花; pufferfish, catfish). In studies with English-speaking children, moreover, this type of cue allows children to more easily learn nouns with the category name embedded in the item label (e.g., oak tree) than those that do not contain explicit category information (e.g., oak) (Gelman, Wilcox, & Clark, 1989). However, the prevalence of words containing such labels differs substantially across these two languages -- they are relatively rare in English, but are highly prevalent in Chinese (Tardif, 2006; X. L. Zhou, Marslen-Wilson, Taft, & Shu, 1999; Y. G. Zhou, 1978). For example, in Chinese, all wheeled vehicles share a common root morpheme (vehicle che1 车 -- e.g., bicycle *zi4xing2che1* 自行车, truck *ka3che1* 卡车, car *jiao4che1* 轿车, bus *gong1gong4qi4che1* 公共汽车, train *huo3che1* 火车). In addition to this level of morphological cuing, over 80% of Chinese characters provide orthographic labels to the category by including a "radical" which labels semantic information (Zhou, 1978; Zhou et al., 1999). For example, although the noun bug *chong2* 虫 is a simple character in its own right, it can also be found as a radical component in many nouns for insects such as fly *cang1ying*苍蝇, butterfly *hu2die2* 蝴蝶, mosquito *wen2zi3* 蚊子, and ant *ma2yi3* 蚂蚁. This kind of orthographic cue can even be found in the oracle bone characters used 3500 years ago (e.g., the radical of water *shui3*  in the characters for river *he2*  and wine *jiu3* ).

In my dissertation, I investigate how differences in noun labelling conventions between English and Chinese nouns influence categorization processes of English and Chinese speakers, and how these similarities and differences are instantiated in the brain. I report evidence that the explicit linguistic cues to category membership in

Chinese nouns facilitate categorization for Chinese speakers, influencing both behavioural measures and neural activity reflecting semantic processing. In four studies, I explore how different types of linguistic cues in English and Chinese nouns (nontransparent, e.g., car, morphologically transparent, e.g., catfish or mei2gui4hua1玫瑰花, and orthographically transparent, e.g., fly cang1ying苍蝇) facilitate the categorization processes of English and Chinese speaking adults, as well as how this difference develops in Chinese speaking children. In the preliminary study (my 619 study), I compared cross-cultural ERP differences between English and Chinese speaking adults when they viewed the same set of pictorial stimuli. In Study 1, I focused on the ERP differences between morphologically transparent items and orthographically transparent items in Chinese speaking adults. In Study 2, I focused on the ERP differences between morphologically transparent items and nontransparent items in English speaking adults. In Study 3, I then examined how these differences might be localized differently in the brain by comparing the differences between English and Chinese speaking adults in an fMRI paradigm when viewing stimuli used in Studies 1 and 2. In Study 4, I focused on the ERP differences between the morphologically transparent items and orthographically transparent items used in Study 2 in 8-9 year old Chinese speaking children. In a integrated model, we should be able to explain how and where the difference between children and adults happened in the brain when they processing exactly the same materials and tasks.

To preview my findings, the results show that English and Chinese speakers activated different neural correlates, despite what appeared to be overall similarities in behaviour (participants respond more quickly/accurately for typical items than atypical items). Specifically, English speakers showed a “typicality effect” in both fMRI and ERP measures, such that atypical objects elicited larger left inferior frontal gyrus (IFG) activation (fMRI) and N300 and N400 components (ERP) than did typical objects. However, none of these typicality effects were apparent for the Chinese participants. Further analysis showed that these differences were mainly between the English nontransparent items and Chinese morphologically transparent items. English morphologically transparent items and Chinese orthographically transparent items actually showed similar activity in the left inferior and medial frontal gyrus and N300 and N400 components. (Liu, et al., in press; Liu et al, submitted). In addition, 8 to 10 year old Chinese-speaking children showed typicality effects for both orthographically and morphologically transparent items, indicating that they could not access the linguistic category information as effectively as adults, and thus did not show the adult pattern of a disappearance of the typicality effect (Liu et al, to be submitted).

These data suggest that cross-linguistic differences in the explicitness of category information have strong effects on the nature of categorization processes performed by the brain. Most importantly, they suggest that speakers of different languages could have different internal category structures, rely on different types of information, and use different neural processes to make category judgments. Moreover, they also suggest that similar structures (e.g., morphological transparency) across languages can result in similar, but not identical patterns of responding.

Chapter II

Preliminary Study (619) – ERP Responses to Identical Pictures across Languages

Method

Participants

Thirty native English speakers in Ann Arbor, MI and 30 native Chinese speakers in Beijing, China, all right-handed undergraduates with normal vision, participated in this study for payment. Four participants in the English group and three participants in the Chinese group were excluded from analysis due to poor behavioral performance (Accuracy < 80% in either Yes or No responses, three English-speaking participants) or too many artifacts in the electroencephalogram data (three Chinese- and one English-speaking participant). The final samples consisted of 27 Chinese speakers (14 females, *Mean* age = 21.81, *SD* = 1.86 years) and 26 English speakers (13 females, *Mean* age = 20.30, *SD* = 1.99) in the behavioral and ERP data analysis.

Development and Design of Stimuli

Because our main interest was in comparing the nature of categorization processes across languages, several steps were taken to ensure cross-linguistic comparability and the validity of our results. First, we avoided the use of linguistic labels as our time-locked stimulus in order to avoid the eliciting ERP components specific to a particular orthography in a word that would differ across languages. Thus, we chose to provide linguistic labels for the categories but pictorial stimuli for the actual items. Moreover, in this Study we chose to present identical pictures to both groups. This necessitated additional pilot testing (see below) to ensure that they were valid and equally typical/atypical across languages. In addition, for this Study it also meant that we had unequal numbers of different types of transparency in the items due to natural variation across languages. This was rectified within each language in Study 1 and Study 2.

In Pilot Study 1, 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?"). This task was used for several reasons. First, it was necessary to verify that the items belonged to the same categories for English and Chinese speakers and that the category-level labels were equally appropriate "substitutions" for both languages, as these stimuli were to be used in subsequent studies. Finally, we wanted to also assess equivalences in labeling and substitutions at the item level for the two language communities. The bilingual and bicultural research team selected a total of 16 categories, each with 3 to 11 items and one or more possible category level labels. By design, both the Chinese and English stimuli contained category items that were *morphologically transparent* and *morphologically non-transparent*. The Chinese stimuli additionally contained items that were *orthographically transparent*.

Twenty native speakers in each location were then asked about the acceptability of replacing each term with another term in that same cluster. The questionnaire was organized by category, but neither the category- nor the item-level labels within a given category were explicitly identified as such and the category-level labels appeared in every possible position (first, second, etc.) across the 16 lists. For example, given the nouns fly *cang1ying1*苍蝇, worm *qiu1ying3* 蚯蚓, bug *chong2zi* 虫子, 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.

Based on the acceptability of replacement judgments from Pilot Study 1, we then eliminated those items or category labels that the majority of participants in one or the other group did not rate as acceptable "replacements" and found corresponding gray-scale photographs for the remaining 2 to 8 items per category.

In Pilot Study 2, twenty-nine English- and twenty-four Chinese- speaking participants were asked to rate the typicality of each picture, given either the category- (e.g., vehicle/车) or item- (e.g., car/轿车) level label, on a 6-point scale, with 1 representing *not at all typical (*完全不典型*)* and 6 representing *extremely typical* (极端典型). These ratings were then used to identify typical and atypical items for each category and to create the final set of ten categories (e.g., vehicle) and twenty objects, half Typical (e.g., car) and half Atypical (e.g., train), used in this Study (see Fig. 1B for labels and pictures and Appendix 1 for the corresponding typicality rating results). The presence of explicit linguistic labels in the item-level labels was reflective of the natural presence/absence of such labels in these two languages and was not possible to control while also ensuring that identical pictures were presented in all conditions. Thus, in English, nine of the ten categories were nontransparent (e.g., car) and only one was morphologically transparent (e.g., writing paper). In Chinese, eight of the ten categories contained morphologically transparent labels, and two were nontransparent. Importantly, there were no differences in the mean ratings of either Typical (English *M*=5.74, *SD*=0.19, Chinese *M*=5.13, *SD*=0.39) or Atypical items across languages (English *M*=3.90, *SD*=0.63, Chinese *M*=*3.70*, *SD*=1.13), with only a main effect of Typicality for these items, *F* (1, 18) = 53.68, *P* < 0.001.

Procedure and Task

English- and Chinese-speaking participants were tested in their native language in either the US or China, respectively. Presentation of the stimuli was controlled with the E-prime program, with participants sitting approximately 20-28 inches away from the screen (resulting approximately 3°of visual angle for words, 1° for crosshair and 10° for pictures) in each location. The presentation procedure can be found in Fig. 1A. Participants first saw either a category-level label (e.g., "VEHICLE" in English or *che1* "车" in Chinese) or an item-level label[i] (e.g., "CAR" in English or *jiao4che1* "轿 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 participant's instructions were to "judge whether or not the picture is an example of the concept represented by the preceding word" ("判断图片是否为其前面的词所代表的概念的一个例子").

A total of 1212 trials were presented, randomly ordered with the constraint that the same words or pictures were not repeated for 3 consecutive trials. The first 12 trials were practice trials. For 500 trials, pictures were preceded by a category-level label and for 700 trials, pictures were preceded by an item-level label. Half of all trials required a *Yes* response (e.g., category-level label VEHICLE *che1*车 followed by a picture of car; or item-level label CAR *jiao4che1* 轿车 followed by a picture of car) and half required a *No* response (e.g., category label VEHICLE *che1* 车 or item level CAR *jiao4che1* 轿车 followed by a picture of eggplant). Half of the 250 "*Yes"* category-level trials (e.g., VEHICLE *che1* 车) were pictures of atypical items (e.g., a train), and half were pictures of typical items (e.g., a car). The experimental session lasted approximately 90 minutes.

EEG Recording

The recording equipment and procedures were nearly identical across the two laboratories, except for the display monitor (12" refresh rate 75 Hz in US and 14" refresh rate 85 Hz in China) and recording software (Neuroscan 4.0 in US and Neuroscan 4.3 in China). The EEG for both sites was recorded using Ag/AgCl electrodes embedded in a nylon mesh cap (21 scalp sites Easy-Cap, Falk Minow Systems Inc., Herrsching-Breitbrunn, Bavaria, Germany) with a left mastoid reference and a forehead ground. An average mastoid reference was derived off-line. The vertical electrooculogram (VEOG) was recorded with electrodes placed above and below the left eye and the horizontal electrooculogram (HEOG) on the outer canthi of both eyes. All interelectrode impedance was maintained below 5 kΩ. The EEG and EOG were amplified using a 0.1 - 100 Hz bandpass filter and continuously sampled at 500 Hz/electrode for off-line analysis with a SynAmps data acquisition system (Neuroscan Labs, Sterling, Virginia, USA). EEG data were corrected for ocular movement artifacts using the Gratton algorithm (Gratton, Coles, & Donchin, 1983). Prior to analysis the data were filtered with a 9-point Chebyshev type II low-pass zero-phase shift digital filter (Matlab 7.0, Mathworks, Inc., Natick, Massachusetts, USA), with a half-amplitude cutoff at 12 Hz.

Results

Behavioural Results

Trials with a response time > 1200 ms or < 200 ms were excluded as outliers (English: 2487 trials, 7.97% of all responses; Chinese: 2203 trials, 6.79% of all responses). A Typicality (Typical *vs.* Atypical) by Language (English *vs.* Chinese) repeated measures ANOVA with Bonferroni corrections for post-hoc analyses was conducted for both accuracy and RT data to explore the effect of Typicality and Language in the *Yes* responses.

As found in numerous previous studies, participants made more errors and responded more slowly when shown pictures of Atypical than Typical members of a category (Accuracy: M = 0.67 and 0.97, F (1, 51) = 270.37, P < 0.001; RT: M = 675.37 and 616.44, F (1, 51) = 142.80, P < 0.001, respectively). However, a significant typicality by language interaction in both the accuracy (P = 0.042) and reaction time data (P = 0.004) indicated that the typicality effect was attenuated for Chinese speakers (Fig. 2A).

ERP Results

The P160, N300 and N400 ERP components were quantified as the positive peak amplitude in the 140-240 (P160) range, and the negative peak amplitude in the 240 - 340 ms (N300) and the 370 - 470 ms (N400) range, respectively. In addition, the LPC component was calculated as the mean positive amplitude in the 500 - 700 ms interval.

All epochs were measured from the onset of the target picture to 800ms later, relative to a 100 ms pre-stimulus baseline.

Based on previous studies of N400 responses and the typicality effect for pictorial stimuli (Fujihara, Nageishi, Koyama, & Nakajima, 1998; Heinze, Muente, & Kutas, 1998; Stuss, Picton, & Cerri, 1988) as well as the scalp topography of the difference waves (Atypical-Typical) for the category-level labels (Fig.3), we focused our analysis on the horizontal line encompassing the bilateral frontal electrodes (F7, F3, Fz, F4, and F8). Voltage data for peak amplitude of the P160, N300 and N400 components from two left-right pairs, F3 - F4 and F7-F8, and mean amplitude for the LPC component for the *Yes* responses were used in a Typicality (Typical *vs.* Atypical) by Side (Left [F3 or F7] *vs.* Right [F4 or F8]) by Language (English *vs.* Chinese) repeated measures ANOVA with Bonferroni corrections for post-hoc analyses (Fig.3). We focus here on the results for the N300 and N400 components.

For both components, main effects of Side and Language were significant for both the F3 - F4 and F7-F8 pairs. The left side elicited a larger negative peak than the right side (F3 - F4 pair $F(1,51) = 27.82$, $P < 0.001$ and F7-F8 pair $F(1,51) = 22.01$, $P < 0.001$ for N300 and $Fs(1,51)=31.70$ and 27.14, $Ps < 0.001$ for N400). In addition, English-speaking participants elicited larger N300s and N400s than Chinese-speaking participants (F3 - F4 pair $F(1,51)=5.39$, $P = 0.024$ and F7-F8 pair $F(1,51) = 5.10$, $P = 0.028$, for N300 and $F(1,51) = 9.89$, $P = 0.003$ and $F(1,51) = 5.39$, $P = 0.024$ for N400). Moreover, the F3 - F4 pair also showed a significant Typicality by Language interaction for both the N300 and N400 components, $F(1,51) = 7.33$, $P=0.009$ and $F(1,51) = 12.33$, $P = 0.001$, respectively, such that the difference between Atypical and Typical items was larger in English- than Chinese-speaking participants. In fact, the typicality effect was almost completely absent from the N300 and N400 components in Chinese speakers. This Typicality by Language interaction was also present for the F7-F8 pair in the N400 component, $F(1,51) = 8.07$, $P = 0.006$, but did not reach significance for the N300.

Interestingly, both English and Chinese speakers showed similar and significant typicality effects in the LPC component such that Atypical items showed larger LPC amplitude at the Fz electrodes than Typical items (Fig. 3). No systematic P160 differences were found between Typical and Atypical items in either language.

Discussion

Despite overall similar behavioral results (albeit attenuated for Chinese speakers) (Fig. 2A), the ERP results in this Study showed dramatic differences in the English and Chinese speakers' processing of typical vs. atypical items. For English speakers, atypical items elicited larger N300 and N400 components than typical items (Fig. 3), consistent with several previous studies in English, German and Japanese (Fujihara et al., 1998; Heinze et al., 1998; Stuss et al., 1988). In contrast, Chinese speakers did not show a typicality effect for either the N300 or N400 components, although they still showed a similar medial frontal LPC (Fig. 3), which might contribute to decision making and evaluative processes or to some sort of post-semantic process such as those found by West and Holcomb (2002) when the goal-related expectations of a pictorially presented story were violated. In our study, the pictorial stimuli were relatively simple, but participants had to decide whether the picture was a "member of the category" with either item- or category-level labels. The fact that the atypical items took longer and also elicited greater positivity across cultures suggests that there were indeed post-lexical processes that may have been related to a final decision process that was more taxing for atypical than for typical items for both groups of speakers. Importantly, when we simplified the decision processes by only including category-level judgment in Studies 1 and 2, this effect either diminished or disappeared altogether.

Chapter III

Study 1 - ERP Responses to Pictures of Chinese nouns

The focus in the preliminary study was on the role of typicality in category-level decisions for speakers of a language that contains category information even in atypical item labels (Chinese) vs. speakers of a language that does not contain category information in the item level (English). The absence of a typicality effect in the N300 and N400 ERP components is not something that has been reported in previous studies and is thus a unique and intriguing aspect of the preliminary study. The Chinese speakers, who have category level information embedded in the names of the items, appear not to rely on typicality during semantic access reflected by the N300 and N400 components. Nonetheless, there were important differences in the numbers of morphologically transparent vs. nontransparent items that were used across the two languages (as a result of their natural frequencies in each language since identical pictures were used) and it is possible that the differences in the language-specific labels for the pictures, rather than the use of morphological transparency as an organizing feature of categories in Chinese vs. English *per se* were responsible for the cross-linguistic differences that we found. Studies 1 and 2 control for these variations in each language separately by adopting additional stimuli. Study 1 controls for the explicitness of linguistic information in Chinese by using pictures with orthographically vs. morphologically transparent labels, and thus asks whether the typicality effect in Chinese speakers is equally absent for orthographically (vs. morphologically) transparent items. Study 2 specifically controls for the explicitness of linguistic information in English by using pictures with morphologically transparent vs. nontransparent labels, and thus asks whether the typicality effect would be similarly reduced in English speakers when provided with pictures that have explicit category information embedded morphologically in their verbal labels.

Method

Participants

Twenty-two (13 females, *Mean* age= 22.10, *SD* = 1.97) native Chinese-speaking undergraduates in Beijing, all right-handed undergraduates with normal vision, participated in this study and received payment.

Stimuli

To generate new categories that included more Chinese items with orthographically transparent items, we first selected 3-4 typical or atypical items for three candidate categories, BIRD, STONE and SHIP (e.g., for BIRD: rooster *gong1ji1*公鸡, duck *ya1z3* 鸭子, penguin *qi3e2* 企鹅, pigeon *ge1zi* 鸽子) and created 2-3 corresponding grayscale pictures for each item. In order to select more typical and atypical items and their corresponding pictures for each category, we conducted two more pilot studies. In Pilot Study 3, we used a naming task in which participants were asked to choose a name to describe each of these new pictures. Twenty native Chinese-speaking undergraduates in Beijing participated and received souvenir pens for participating. Using these results, we then removed pictures that received < 60% naming accuracy to the intended item name. (This is a conservative criterion since we were looking for exact hits and many non-hits were very close in meaning with similar orthographic or morphological structures as the intended names). In Pilot Study 4, we asked participants to rate typicality for all remaining item pictures using the same scale (1-6) as in Pilot Study 2. Twenty-three native Chinese-speaking undergraduates in Beijing participated and received souvenir pens. Based on these results, we then discarded items that had intermediate or below-threshold typicality ratings and kept the two items for each category with the lowest and highest typicality ratings (Appendix 2). Finally, we selected six pictures for these three new orthographical categories: BIRD *niao3*鸟, STONE *shi2tou2*石头, SHIP *chuan2* 船 (Fig. 1C, Appendix. 2). In our final stimulus set, we had five morphologically transparent and five orthographically transparent categories for Study 2, each with one Typical and one Atypical picture (Fig. 1C).

Procedure and Task

The procedure, apparatus and task of Study 1 were the same as for the Preliminary Study except that only category-level trials were included so that we could reduce the testing time and task difficulty. For practical reasons, we also changed the ERP recording system from Neuroscan with 21 scalp sites (Easy-Cap) to EGI with a 128-channel Geodesic sensor Net and EGI NetStation 4.1.

A total of 412 trials were presented in pseudo-random order to each participant. The first 12 trials were practice trials. As with the Preliminary Study, half of all trials were correct that required a *Yes* response (e.g., category label VEHICLE *che1* 车, followed by a picture of car) and half were wrong that required a *No* response (e.g., label VEHICLE *che1* 车 followed by a picture of eggplant). Among the 200 *Yes* trials, half were Typical and half were Atypical, and this was crossed with Label type (Morphological *vs.* Orthographical), yielding 50 trials for each condition (e.g., Typical Morphological items). The experimental session lasted approximately 25-30 minutes.

EEG Recording

The electroencephalogram (EEG) was recorded using a 128-electrode Geodesic Sensor Net. The EEG signal was amplified using a 0.01 – 100 Hz bandpass and digitized at 500 Hz. The electro-oculogram (EOG) was monitored with 6 electrodes placed bilaterally in the external canthi (128 and 125), supraorbital (26 and 8) and infraorbital (127 and 126) regions. Impedances for each electrode were measured prior to recording and kept below 50 kΩ during testing. Recording in every electrode was vertex-referenced. In order to compare data across studies, we also conducted a second set of analyses with an average mastoid referencing procedure. Results from both methods were nearly identical, but the average of all electrodes method will be presented as our primary findings since this is the method most appropriate for the EGI system and is a better representation of a true zero for the Geodesic Sensor Net (Junghofer, Elbert, Tucker, & Braun, 1999). Deviations across referencing methods will be noted where relevant. The 100 ms preceding the target served as baseline. Data were recorded and processed using Net Station 4.1 (EGI software).

After acquistion, the data were lowpass filtered below 20 Hz. The continuous EEG was segmented into an epoch starting at 100 ms before the onset of the stimulus and lasting until 800 ms after stimulus onset. Segmented files were scanned for artifacts with the Artifact Detection toolbox in NetStation 4.1 using a threshold of 70 μV for excessive muscular activity, eye blinks and eye movements. Segments containing eye blinks or movements as well as segments with more than 20 bad electrodes were rejected. Within each segment, electrodes with either an average amplitude of greater than 200 μV or difference average amplitude of 100 μV were also discarded from further processing. Finally, particular electrodes were rejected if they contained artifacts of any kind in more than 50% of the segments. Artifact-free segments for correct responses were averaged separately for *Yes* and *No* trials over the 800-ms epoch across subjects and re-referenced against the average of all electrodes.

Results

Behavioural Results

Trials with a response time > 1200 ms or < 200 ms were cut off as outliers (304 trials, 3.6% of all responses). Fig. 2B shows the accuracy and RT data to pictures with Morphologically *vs.* Orthographically transparent labels and Atypical *vs.* Typical items in *Yes* responses. A Typicality (Typical *vs.* Atypical) by Label type (Morphological *vs.* Orthographic) ANOVA found that the main effects of Typicality and Label type were significant specifically for both RT and accuracy data. Participants made more errors, $F(1,21)=16.12$, $P<0.001$ and responded more slowly, $F(1,21)=37.23$, $P<0.001$ for Atypical than Typical items. They also made more errors,

$F(1,21)=15.86$, $p<0.001$ and responded more slowly, $F(1,21)=17.09$, $P<0.001$ for Orthographically transparent items than Morphologically transparent items. A significant Typicality by Label type interaction was also found in the accuracy data, $F(1,21)=13.39$, $P=0.001$, such that the typicality effect was larger for the Orthographically transparent items than Morphologically transparent items.

ERP Results

The P160, N300 and N400 ERP components were quantified as the negative or positive peak amplitude in the 130-190ms (P160), 240-340 ms (N300) and the 370-470 ms (N400) range, respectively. In addition, the LPC component was calculated as the mean amplitude in the 500-700 ms interval. All epochs were measured following the onset of the target picture, relative to a 100 ms pre-stimulus baseline.

As in the Preliminary Study, we focused our analysis on the horizontal line encompassing the bilateral frontal electrodes (F7, F5, FCz, F6, and F8, corresponding to electrodes 34, 28, 6, 123, 122 in the Geodesic sensor Net, respectively) (EGI software). Voltage data for peak amplitude of the P160, N300 and N400 components from two left-right pairs, F5 - F6 and F7 - F8 for the *Yes* responses were used in a Typicality (Typical *vs.* Atypical) by Side (Left [F5 or F7] *vs.* Right [F6 or F8]) by Label type (Morphological *vs.* Orthographic) repeated measures ANOVA with Bonferroni corrections for post-hoc analyses (Fig.4).

As can be seen in Fig. 4, no significant typicality effects were found when combining both types of stimuli for both N300 and N400 components ($F(1,21)=0.183$ and 0.371, $Ps > 0.549$), which repeated our finding in the Preliminary Study. However, the F5-F6 pair showed a marginally significant Typicality by Label type interaction for the N400 component, $F(1,21)=2.93$, $P=0.10$, such that orthographically transparent items indeed showed left-lateralized typicality effects for both the N300 and N400 components (Fig.4), whereas morphologically transparent items showed no significant differences between Typical and Atypical items. A further Typicality (Typical *vs.* Atypical) by Side (Left [F5 or F7] *vs.* Right [F6 or F8]) ANOVA for Orthographically transparent items only revealed that Atypical items elicited a larger N400 (F5-F6 pair $F(1,21)=4.92$, $P=0.038$ and F7-F8 pair $F(1,21) = 5.24$, $P = 0.033$) than the Typical items.

The LPC component showed no significant main effects of Typicality or Typicality by Label type interaction, which was true also for the average mastoid referencing results.

Discussion

Interestingly, and consistent with the findings from the Preliminary Study, native Chinese speakers did not show a typicality effect in the N300 and N400 components for morphologically transparent items. In contrast, and parallel to the cross-linguistic findings when we compared English *vs.* Chinese speakers, orthographically transparent items revealed larger typicality effects in both behavioral and ERP results than morphologically transparent items. For Chinese speakers, orthographically transparent items showed a significant typicality effect for both the N300 and N400 components in the left frontal electrodes (Fig. 4).

These results suggest that even when category judgments for pictures are used, different label types influence Chinese speakers' use of typicality to make these judgments. Although orthographically transparent nouns provide category information in Chinese, it appears to be less accessible than the information provided by morphologically transparent nouns. Since the orthographic information embedded in Chinese radicals is not pronounced and does not also provide phonological information when the label of an item is accessed, the category information provided by orthographically transparent labels may thus be more implicit, and also does not become available for use until Chinese speakers become fluent readers (i.e., not from the beginning of productive language). In addition, fMRI studies on Chinese character processing have found that orthographical information in Chinese characters may also require additional orthographic-to-semantic mappings in order to be accessed (C. Liu et al., 2008; Siok, Jin, Fletcher, & Tan, 2003; Siok, Perfetti, Jin, & Tan, 2004; Tan, Laird, Li, & Fox, 2005) and thus may need more semantic processing to process than morphologically transparent items. However, it is not clear whether orthographic information might still confer an additional advantage, albeit very small, relative to a completely nontransparent item. This was not tested and would perhaps be undetectable in the current design, but is worthy of future study, particularly given that the differences between the morphologically and orthographically transparent items in the present study were significant, but relatively small compared to the differences between English and Chinese in the previous study.

These results also provide data on how to build an integrated model of categorization. None of the previous models considerate the linguistic factors in categorization. However, our results on Chinese speaking adults clearly showed that different linguistic label types will have different impact for the categorization processing performed in the language speakers' brain. Thus, an integrated model should be able to explain how and why this influence happens.

Nonetheless, this label type difference in Chinese is similar to the cross-cultural difference found in the Preliminary study. Just as nontransparent English nouns that provide no category information need more semantic processing in order to make category judgments, orthographically transparent Chinese items that provide less salient linguistic category information need more semantic processing than morphologically transparent items. Relying on typicality, for less transparent items, is thus a useful way to reduce the amount of semantic processing required. Because morphological transparency, in Chinese, is both a regular feature of the language and it provides explicit and solid linguistic cues to category membership, typicality is not needed for initial category access and thus morphologically transparent items in Chinese do not show a typicality effect.

Chapter IV

Study 2 - ERP Responses to Pictures of English nouns

Our next question then is whether or not morphologically transparent items in English (e.g., cat<u>fish</u>) facilitate categorization processes in English speakers. Although this type of transparency does occur in English, it is not as productive and regular as it is in Chinese (Tardif, 2006; X. L. Zhou et al., 1999). Thus we are able to use English to distinguish between one of two possibilities. First, our results in Chinese may simply be an immediate effect of the relation between a picture's label and its category; if so, we might expect that English speakers can also rely on the explicit category information provided by morphologically transparent items and thus show a reduction in the typicality effect just like Chinese speakers when pictures of morphologically transparent items are used in a category judgment task. However, if it is not only an immediate effect of label type, but also the conventions of a language that play a role, then English speakers might not be able to extract such linguistic information as efficiently as Chinese speakers, who have presumably used such linguistic cues in their implicit processing of language since they first began to understand and produce words (Tardif, 2006).

Method

Participants

Thirty-three native English speakers in Ann Arbor, MI, all right-handed undergraduates with normal vision, participated for course credit. Eight participants were excluded from further analysis, four for poor behavioral performance (Accuracy < 80% in either *Yes* or *No* responses) and four for too many eye-blinks or artifacts in the electroencephalogram data. The final sample consisted of 25 participants (8 females, *M* age = 19.12, *SD* = 1.04).

Stimuli

To generate new categories that included more English items with morphologically transparent items, we followed the procedures of Pilot Studies 3 and 4. First we selected 3-4 typical or atypical items for eight candidate categories, PHONE, POOL, BAG, BOOK, BALL, CHAIR, STATION, and PAPER (e.g., BALL: basketball, football, soccer ball, and baseball) and produced corresponding grayscale pictures for each item. In Pilot Study 5, twenty-nine native English-speaking undergraduates provided labels for and rated the typicality of all pictures on a scale from 1 to 6. Based on these results, we then discarded items that had poor label agreements and/or intermediate typicality ratings and kept the two items for each category with the lowest and highest typicality ratings (Appendix 3).

For Study 2, we selected four of these morphologically transparent categories (BALL, BOOK, CHAIR, PHONE), each with a typical and an atypical item, and combined them with the one morphologically transparent category (PAPER) and five nontransparent categories used in Study 1, to produce a total of five morphologically transparent and five nontransparent categories (Fig. 1D, Appendix. 3).

Procedure and Task

The procedure, apparatus and task of Study 2 were the same as for Study 1.

A total of 812 trials were presented pseudo randomly to each participant as in Study 1. The first 12 trials were practice trials. Half of all trials required a *Yes* response and half required a *No* response. Among the 400 "*Yes"* trials, half were Typical and half were Atypical, and this was crossed with Label type (Morphological *vs.* Nontransparent), yielding 100 trials for each condition (e.g., Typical Morphological items). The experimental session lasted approximately 45-50 minutes.

EEG Recording

The electroencephalogram (EEG) recording and analyses procedures and equipment for Study 2 were identical to Study 1. The study was conducted in the U.S. with identical EGI equipment and software as the laboratory in Beijing.

Results

Behavioural Results

Trials with a response time > 1200 ms or < 200 ms were excluded as outliers (774 trials, 3.9% of all responses). As with Study 1, a Label type (Morphological *vs.* Nontransparent) by Typicality (Typical *vs.* Atypical) repeated measures ANOVA revealed that participants made more errors, $F(1,24)=6.07$, $P=0.021$ and responded more slowly, $F(1,24)=13.94$, $P=0.001$, for Atypical items than Typical items. In addition, English speakers were slightly faster to categorize exemplars that contained Morphologically transparent items than those that did not, $F(1,24)=6.20$, $P=0.020$, thus suggesting that even for English, morphological transparency can convey a slight advantage during categorization. However, unlike the cross-linguistic comparisons in Study 1, no interactions were observed between typicality and the morphological transparency conditions for either the RT or accuracy data (Fig. 2C).

ERP Results

The P160, N300, N400 and LPC ERP components were quantified the same way as in Study 1 and analyzed using the same software.

As in the Preliminary study and Study 1, we focused our analysis on the horizontal line encompassing the bilateral frontal electrodes (F7, F5, Fcz, F6, and F8, corresponding to 34, 28, 6, 123, 122 in the Geodesic sensor Net, respectively). Voltage data for peak amplitude of the P160, N300 and N400 components from two left-right pairs, F5-F6 and F7-F8 for the "*Yes*" responses were used in a Typicality (Typical *vs.* Atypical) by Side (Left [F5 or F7] *vs.* Right [F6 or F8]) by Label type (Morphological *vs.* Nontransparent) repeated measures ANOVA with Bonferroni corrections for post-hoc analyses (Fig. 5). As with Study 1, we focus here on the N300 and N400 components. As can be seen in Fig. 5, significant typicality effects were found for both types of stimuli, but laterality interacted with the type of information provided in the stimulus label - the typicality effect for the nontransparent items appeared in left-side scalp electrodes whereas the typicality effect for the morphologically transparent items appeared in right-side scalp electrodes and this was true for both the N300 and N400 components.

Atypical items elicited a larger N300 (F5-F6 pair $F(1,24)=15.21$ and F7-F8 pair $F(1,24) = 15.22$, $P < 0.001$) and N400 (F5-F6 pair $F(1,24) = 8.46$, $P < 0.01$ and F7-F8 pair $F(1,24) = 14.51$, $P < 0.001$) than the Typical items. In addition, both the F5-F6 and the F7-F8 pairs showed significant Typicality by Label type by Side interactions for both the N300 ($F(1,24)=9.48$ and 8.82, $Ps < 0.01$) and N400 ($F(1,24)=5.45$ and 5.44, $Ps < 0.05$) components, such that the Nontransparent items showed a significant typicality effect only on the left side electrodes, whereas Morphologically transparent items showed a significant typicality effect only on the right side electrodes(Fig.5).

The LPC component showed no significant main effects of Typicality or Typicality by Label type interaction, which was true also for the average mastoid referencing results.

Discussion

Although the behavioral results in Study 2 were almost identical to those for the English speakers in the Preliminary Study, the ERP results revealed some interesting differences in the typicality effect for nontransparent vs. morphologically transparent items for English speakers. Nontransparent atypical items elicited larger N300 and N400 components than typical items only in the left frontal electrodes, whereas for the morphologically transparent nouns, this difference was apparent in the right frontal electrodes, as shown in Fig. 5. The laterality of the typicality effect for English nouns is intriguing, but it is not clear from these data alone whether it truly reflects differences in right vs. left hemisphere processing of morphologically transparent vs. nontransparent nouns, or whether there are some other differences contributing to the appearance of this effect at different electrode sites. Nonetheless, as has been found in the Preliminary Study and Study 1, the increased activation for atypical nouns indicates increased semantic processing (Fujihara et al., 1998; Heinze et al., 1998; Poldrack et al., 1999; Stuss et al., 1988). In English, this is particularly prominent for pictures of nouns that do not contain any linguistic cues to category information, which is the predominant pattern in English. In contrast, when pictures of items that have morphologically transparent cues to the category are provided, even English speakers appear to make use of these cues to facilitate semantic access and category judgments, as evidenced by the decrease in RTs for morphologically transparent items. However, since most English nouns are nontransparent rather than linguistically transparent like Chinese nouns, providing pictures of morphologically transparent items is not enough to allow English speakers to circumvent the use of typicality as an aid to categorization. As a result, additional executive attention load (Corbetta & Shulman, 2002; Han et al., 2004) might be involved for English speakers' analysis of the morphologically transparent category information. We interpret the relatively greater activation for the atypical morphologically transparent nouns at right scalp electrodes to be a result of this additional processing load for English speakers. These data alone, however, do not allow us to clarify whether the nature of the typicality effect is identical for morphologically transparent vs. nontransparent items in English.

Again, just like those ERP differences we found between the orthographically transparent items and morphologically transparent items in Chinese speaking adults, these ERP differences between morphologically transparent items and nontransparent items in English speaking adults could not be explained by any of the previous categorization processing models. We need an integrated model to explain why and how the differences of linguistic labels influence categorization in the brain.

Chapter V

Study 3 - fMRI Responses to Category Judgements across Languages

Although we already revealed how different English and Chinese nouns differ in providing category information in Study 1-2, we still did not reveal exactly where these differences happen in the brain, mainly due to the poor spatial resolution of the ERP technique. In particular, how English and Chinese speakers are similar and different in performing categorization judgments across the three categorization systems (E. E. Smith & Grossman, 2008) is still unknown. Although there has been extensive psychological and neuroscience research on similarities and differences in the ways in which English and Chinese words and characters are read and processed (e.g., Chan et al., 2009; Chou et al., 2009; Liu et al., 2009) and some research on the processing of tones and grammar of Chinese (see Zhou et al, 2009 for an overview), very little research has been conducted on the ways in which English and Chinese speakers process other aspects of language such as the nature of the words that we use to label and categorize the world and how this is related to other aspects of cognition (but see Liu et al., in press; Mok et al., 2009; Tan et al., 2008). In one series of studies, Siok and colleagues have found that the ways in which color terms are encoded in a language affect color perception for visual stimuli presented in the right visual field as well as for which areas of the brain are activated when "within" vs. "between" color label discriminations are made. These findings impact our understanding of the role of top-down linguistic processing on visual perception in the brain, as well as how colors may be processed by the brain for speakers of different languages (Siok et al., 2009). These studies, however, do not ask whether there are fundamental differences in whether and how speakers of different languages make semantic judgments per se.

In Study 3, we reinvestigated the same research questions in as Studies 1-2 using the fMRI technique. Compared with ERP, fMRI can provide much more precise spatial resolution and can therefore be a useful tool in understanding which brain structures and ultimately, which types of information are being processed during categorization judgment.

Method

Participants

Twenty native 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 paid RMB100 (approximately US$12). Three participants were excluded from further analysis, two Chinese speakers for poor behavioral performance and one English speaker for uncorrectable head movement (> 4mm) 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.

Stimuli

Twenty-eight greyscale object pictures for 14 categories were selected from our previous study (C. Liu et al., in press) (Fig.1 C and D). To ensure consistency across categories and languages, every participant completed a typicality rating survey before the scan by first naming the item 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？). All item pictures received >80% naming accuracy at the item-level from both English and Chinese speakers. The typicality rating results for the six categories compared cross-linguistically (Appendix 4) and ten categories in English (Appendix5) and Chinese (Appendix 6) revealed no significant interactions between typicality and other factors of interest such as language or label type, thus ensuring the comparability of the stimuli across languages and conditions in the current study.

Procedure and Task

The procedure and task can be found in Fig. 1 A. During the fMRI scan, a total of 400 trials, 20 presentations for each of the 20 pictures, were presented in random order to each participant in four separate runs with 100 trials each. The inter-trial interval was jittered at 500, 2000, 3500, 5000, and 6500ms 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 (Fig. 1 C and D, bottom six categories) in English and Chinese, yielding 60 trials for each language. The whole experimental session lasted approximately 1 hour with 25-30 minutes for fMRI data acquistion.

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).

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/. 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 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 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 128s, and a low-pass filter was achieved by convolution with the standard SPM hemodynamic response function (HRF). 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 involved a two sample *t*-test with an uncorrected voxelwise threshold of $P < 0.001$ and clusterwise threshold of $P < 0.05$ (K >10) (Canli et al., 2005) with False Discovery Rate (FDR) corrected for multiple comparisons using the small volume correction (SVC). The within-linguistic contrasts involved one sample *t*-tests a revealed weaker brain activation in general, thus a reduced threshold of uncorrected voxelwise $P < 0.005$ was set with a FDR corrected clusterwise $P < 0.05$ ($K > 20$ 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. Based on the frontal and parietal regions identified in various previous categorization studies (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) and the results of the Yes *vs.* No contrast and cross-linguistic comparison of Atypical and Typical items (Fig. 7, 8, Tables 2, 3), small-volume correction (SVC) was done separately using ROIs within the left BA 46 and 47, the right BA46 and 47, 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. 9) (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.

Results

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).

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 for the six cross-linguistically identical categories to explore the effect of Typicality and Language in the "yes" responses (Fig.6 A). Overall, 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) =324.71 , P <0.001, and made more errors and responded more slowly for Atypical items than Typical items [Accuracy: M = 0.96 and 0.99, F (1, 34) = 10.70, P = 0.002; RT: M = 569.14 and 527.67, F (1, 34) = 53.52, P < 0.001, respectively], as shown in Fig. 6 A, which replicated 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 = 588.05 and 508.76, F(1, 34) = 8.90, P = 0.005) and generally gave higher ratings (Typical M = 5.71, Atypical M = 3.63) than English speakers (Typical M = 5.21, Atypical M = 2.80) for both types of items, with no significant Typicality by Language interaction in the ratings or in the accuracy data. However, there was a significant Language by Typicality interaction in the RT data, F (1, 34) = 5.17, P = 0.029, such that English speakers appeared to show a weaker typicality effect (M = 523.05 ms for atypical vs. 494.47 ms for typical) than Chinese speakers (M = 615.23 ms for atypical vs. 560.87 ms for typical), suggesting that English and Chinese speakers differed in the extent to which they showed the Typicality effect. Given that these results were not consistent across the three measures, and were found in previous studies with the same stimuli (e.g., Liu et al., in press), it is not clear whether it was a true interaction or whether it was simply a magnification of the effect given the slower RT of the Chinese participants. Overall, even considering the apparently stronger typicality effect in the RT for Chinese participants for 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.6 B). 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 they made more errors and responded more slowly for Atypical items 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 Morphological labels than Nontransparent labels (M = 512.30 ms and 500.34 ms, F(1, 17) = 8.78, P = 0.009). However, no interactions between Typicality and Label Type were found for any of the measures and English speakers generally gave similar ratings for Morphologically cued labels (Typical M = 5.19, Atypical M = 2.97) as they did for Nontransparent labels (Typical M = 5.24, Atypical M = 2.81).

For the Chinese speakers, 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 items 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 Morphological labels (Typical M = 5.76, Atypical M = 3.62) than Orthographical labels (Typical M = 5.42, Atypical M = 3.19), and responded more slowly for Orthographic labels than Morphological 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 Figure 6 C.

Imaging Results

In contrast to the behavioral results, 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. 7, Table 1). In contrast, Chinese speakers showed no differences between typical and atypical items in these regions (Fig. 7, Table 1). Moreover, a cross-linguistic contrast between English and Chinese speakers in typical and atypical items further revealed that the differential frontal activity between typical and atypical items mainly comes from the atypical items rather than typical items (Fig. 7, Table 2).

And yet, when contrasting responses to items belonging to the category (*Yes* responses) versus those not belonging (*No* responses) (Fig. 8A, Table 3), 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.8A, Table 3). This indicates that the absence of frontal activation differences in the typicality effect is not some quirk of sampling or how Chinese speakers performed the task. Moreover, these results echo findings from Studies 1 and 2 using ERP paradigm

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. 6), 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 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. 8B, 9; Table 4).

Discussion

As far as we know, no previous fMRI study has explored the localization of the "typicality effect" involved in semantic categorization. Nonetheless, our results are consistent with other studies involving semantic categorization. Specifically, 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 et al., 2007; Grossman et al., 2002; Jiang et al., 2007; Koenig et al., 2005; Myers, 2007; Reber et al., 2003). The left inferior frontal gyrus, in particular, has been identified as a region which contributes greatly to semantic and lexical access for English speakers (Bookheimer, 2002; Hagoort, Hald, Bastiaansen, & Petersson, 2004; Vandenberghe, Price, Wise, Josephs, & Frackowiak, 1996) as well as bilinguals (Rodriguez-Fornells, Rotte, Heinze, Nosselt, & Munte, 2002) and monolingual Chinese speakers (C. Liu et al., 2008; Siok et al., 2004). For out-of-category items (*No* responses) than in-category items (*Yes* responses), our study echoes these findings of greater left inferior frontal gyrus activation for both the English- and Chinese speaking participants. Moreover, in English, atypical items showed larger left IFG activity than typical items both for nontransparent and morphologically transparent items. In Chinese, this pattern was also true for orthographically transparent items (Fig. 8B, 9; Table 4). These results further implicate the role of semantic processing in the generation of the typicality effect.

In addition to the left IFG, both groups of participants also showed a typicality effect in the bilateral SFG or 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 et al., 2005; Feredoes et al., 2006) rather than the parietal areas that were also active during this version of the task which involved integrating visual with semantic information.

Of great interest, however, is whether participants in the two cultures engaged in the same type of semantic access and/or whether they engaged in additional semantic processing of the pictorial stimuli, given the differences in their lexical labels (e.g., car/*jiao4*<u>*che1*</u>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 a typicality effect for the Nontransparent Atypical items in the IFG (bilaterally) and MFG (primarily left hemisphere), 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 the category name in the common label for morphologically transparent items (e.g., *jiao4*<u>*che1*</u>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 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), as would be expected for implicit rule-based categorization (Grossman et al., 2002). Interestingly, however, when English participants were given a morphological cue, it did not result in the same type of semantic processing bypass, since 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. 8B). Clearly, though, providing a morphological cue to category membership in English is not as effective as it is in Chinese, a language which has had a long tradition of morphological cues to category for thousands of years across a large number of semantic categories. Interestingly, because the radicals in the 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. This finding is echoed by another 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 for Chinese, was found to be critically related to semantic, but not phonological, processing in Chinese speakers (Liu et al., 2009).

In order to account for the above patterns of brain results from Study 1 to Study 3, I proposed that a revision to the Farah and McClelland ‘s model of semantic memory that might accout for the brain data we obtained here. As can be seen in Fig.11, after viewing the category name and object picture sequentially, the verbal words and visual pictures are first encoded and recognized by perception systems in the temporal and occipital cortex, respectively, which happens about 150 to 200 ms, as represented by the N1 ERP component. These perceptual inputs then enter the working memory system of either the visual buffer or the phonological buffer (Baddeley and Hitch, 1974; Baddeley, 2000), in which the name of both the category and the object are decoded and compared in the phonological buffer whereas the physical properties of both the category and the object are decoded and compared in the visual buffer. This visual and phonological information stored in working memory then enters long term memory, where a comparison is made between the structure and those category properties stored in the conceptual knowledge. Based on E. E. Smith & Grossman (2008), these lexical and semantic processes in working and long term memory were processed in the frontal cortex, especially inferior frontal gyrus and middle frontal gyrus, and happen in a window from 250 to 500 ms, as represented by the N300 and N400 component. Finally, comparing results from long term memory are sent to the executive system in the superior and medial frontal cortex(Corbetta & Shulman, 2002; Grinband, Hirsch, & Ferrera, 2006), which make the final decision about the category judgment, represented by those later components starting from 600 ms (Stuss et al., 1988;West and Holcomb, 2002).

In this model, the differences between different linguistic label types influence both lexical access and verbal encoding in working memory, whereas the difference between the language labelling conventions and speakers' experience with it influence the comparing and accessing of this information in long term memory. Thus, morphologically transparent items, which share overlapping verbal information as the category name, regardless of whether they are typical or atypical, when being encoded in the phonological buffer of the working memory, could bypass the accessing and comparison process in long term memory (as indicated by the arrow in Fig. 13), resulting in an attenuated typicality effect in the N300 and N400 components.

For orthographically transparent items that are unpronounceable, such bypassing will require an additional visual decoding processing of the semantic radical in the item names, resulting in an attenuated typicality effect in the brain for Chinese speakers, as compared with nontransparent items in English speakers.

Chapter VI

Study 4 - Developmental Trajectory of the Influence of Chinese nouns on Categorization

Based on the adult behavioural and neuroimaging results, I conducted one more study to further explore how these behavioural and brain differences are developed in Chinese speaking children. The main finding from previous studies is the disappearance of the typicality effect in the brain (either N300 and N400 ERP components or left middle/inferior frontal activation in fMRI) for Chinese morphologically transparent items (Fig. 3, Fig.4, Fig.7). We interpreted this finding to be a direct influence of noun labelling conventions in Chinese for categorization processing, such that explicit linguistic cues to category membership could provide a "shortcut" to the category membership judgement and thus eliminate the typicality effect. However, all of these findings were from Chinese speaking adults. How linguistic labels influence categorization processing in Chinese speaking children is still unclear. Although the most commonly used adult behavioural indexes of accuracy and reaction time are not accurate in reflecting children's behavioural responses, other measurements such as category learning have identified a similar "typicality effect" in children. For instance, Mervis & Pani (1980) taught children labels for new object categories. Some children were taught the categories based on exposure to good (or typical) exemplars, and others were taught the categories based on exposure to poor (or atypical) exemplars. After training, they were presented with objects they had not seen before, some of which were members of the category and others which were not. As predicted, children who had been exposed to the good (or typical) exemplars were better able to learn the category terms than those who had been exposed to the poor (or atypical) exemplars. A further experiment showed that even when exposed to a range of exemplars, children will tend to learn the good (or typical) exemplars of a category more readily than the poor (or atypical) exemplars (Mervis & Pani, 1980). They argued that the good (or typical) examples provide the most accurate basis for generalization to new instances because they are maximally similar to members of their own category and minimally similar to members of other categories. In fact, children tend to learn typical instances of a category before atypical ones, perhaps because parents tend to teach the typical ones first (Anglin, 1977).

Several studies have found that language plays an important role in the way that object categories are organized and structured in children's minds (Martinez & Shatz, 1996; Yoshida & Smith, 2003). For example, one study found that after naming a novel object, children are much more likely to categorize it based on conceptual features (taxonomically) than perceptual features (superficially) (Gelman & Markman, 1987). Learning new items in a category was also easier when children were presented with compound words with the category name (i.e., oak tree) versus those that were not (i.e., oak) (Gelman, Wilcox, & Clark, 1989). Similar results have also been found in Japanese children, such that Japanese 2-year-olds were better able to map a novel word for a familiar animal to a subordinate category when it was presented as a compound noun (e.g., X-penguin (penguin))(Imai & Haryu, 2001). These studies indicate that children are able to utilize a linguistic cue that is available in their language to modify a default interpretation. In addition, naming objects in a category even encourages learning of other objects from the same category in 12-month-old infants (Waxman & Markow, 1995).

Studies on Chinese speaking children have found that these noun labelling conventions also predominate in children's vocabularies. Tardif and colleagues (Tardif, 1996; Tardif, Fletcher, Liang, & Zhang, 2002) conducted studies that involved adapting the MacArthur-Bates Communicative Development Inventory (CDI) in Chinese and Cantonese speaking children and noticed an interesting phenomenon when attempting translate animal nouns, household objects, toys, and everyday items. Specifically, many distinct English nouns had equivalents that were compound words consisting of a prefix and a common category term in Chinese. For example, in English, rooster and hen are both types of chickens. Nevertheless, in Chinese, all 3 of these nouns share a common category label, chicken *ji1*, and hen and rooster are productive variations with the category labels female *mu3* and male *gong1*. Moreover, these male-female prefixes and other labels were used over and over again for other animals. In contrast, the English nouns "mare" and "cow" have no obvious morphological relations to "hen," despite the fact that they are all females (Tardif, 2006).

However, how Chinese speaking children differ in categorization from English children is still unclear. Chiu (1972) reported that in a categorization task, Chinese children were more likely to categorize objects based on shared relationships, whereas American children were more likely to categorize objects based on similarity. Similar results have also been found in adults, such that when asked to look at sets of three pictures and to decide which two pictures of each set best belonged together, Chinese adults were equally likely to group items together if they shared a relationship (e.g., tire-car) and if they shared a category (e.g., bus-car), whereas American adults were more likely to group items together if they belonged to the same category (Unsworth, Sears, & Pexman, 2005). However, these results usually are interpreted from social psychological viewpoints, such that westerners are presumed to engage in more context-independent and analytic perceptual processes whereas Asians tend to engage in more context-dependent and holistic perceptual processes (for a review, see (Nisbett & Miyamoto, 2005). Very little research has been done from the viewpoint of language differences even thought these very findings could also be explained by the linguistic cues embedded in Chinese nouns (e.g., both tires and cars in Chinese contain the morpheme for vehicle, *Che1* in their names, although tire *lunzi* contains it in the orthographic information, and car, *qi4che1* as a morphological component of the spoken word and syllable).

ERP studies on children's categorization have found similar components to those found in adults. In a visual category study (Quinn, Westerlund, & Nelson, 2006), 6 month-old infants viewed cat images during familiarization, followed by novel cat images interspersed with novel dog images. Specifically, a negative central (NC) component peaking at around 500ms from left-central electrodes was found exclusively for novel dog images, which is quite similar to the N400 component found for out-of-category or atypical objects in adults (Fujihara et al., 1998; Heinze et al., 1998). In a cross-sectional study with 48 children (aged from 7 to 15 years old) and 14 adults, Batty and Taylor (2002) found that children beginning at even 7 years of age preformed the animal/nonanimal visual categorization task with natural images very similarly to adults, with responoses as accurate (96%) as adults with only 80ms slower reaction times. In this study, however, the developmental trajectory of different ERP components was also investigated for another index of categorization, the N2 latency, which reached adult levels at 9 years, and the N2 amplitude changes start at 12 years of age.

Another common ERP component found in children is late slow waves (LSWs). LSWs have been found to be associated with processes of working memory systems. For example, different LSW scalp topographies have been found as a function of the type of materials being operated on working memory, such that phological memory operations elicted the largest LSW over the left hemisphere scalp sites, whereas visual feature memory operations elicted largest LSW over the right hemisphere scalp sites (Barrett et al, 1988; 1989). In addition, Uhl et al (1990) found that slow waves were largest over parietal scalp areas when subjects imagined a diagram but largest over temporal and occiptal scalp areas when they imagined faces and colors. Moreover, two previous ERP studies have investigated false-belief reasoning in adults (Liu et al, 2004; Sabbagh & Taylor, 2000); with both studies finding that false-belief reasoning was associated with a LSW over left-frontal regions. Such a belief-reasoning related LSW has also been identified in 4 to 6 year old children while performing a false-belief task with animated vignettes, such that children who could correctly reason about the characters' beliefs showed larger LSWs in the left-frontal regions, whereas children who failed false-belief questions did not (Liu et al, 2009).

Nevertheless, to my knowledge, there is no ERP study that yet investigates Chinese children’s categorization. Given these previous findings on the influence of language on the categorization processing of Chinese speaking adults, my research question is “How precisely might these effects in adults arise over the course of earlier lives/experience?” In particular, because children (e.g., 8-9 years old, Grade 2 or 3) do not have as much experience as adults in both verbal and written language, we might expect much less influence from Chinese language on their categorization processing, as reflected by different neural signatures (e.g., ERP) patterns for the typicality effect. Most interestingly, how could the ERP pattern of Chinese speaking children be compared with that of Chinese speaking adults and English speaking adults to tell us more about the origin and meaning of those effects we found in previous adults studies?

One of the most important findings in the studies of Chinese speaking adults Study 1 and 3 is the difference between morphologically transparent items and orthographically transparent items. For Chinese speaking adults, morphologically transparent items showed no reliable typicality effect in the N300 and N400 components, whereas orthographically transparent items still showed a significant typicality effect for both the N300 and N400 components in the left frontal region (Fig. 4). I proposed that this is because the orthographic information embedded in written Chinese radicals is not pronounced and does not therefore provide phonological information when an item might be accessed, the category information provided may be more implicit, later acquired (i.e., once children are able to read, not from the beginning of productive language), and harder to access compared with morphologically transparent items, which have phonological as well as morphological cues to category membership that is embedded in the name of the item.

One way to verify this hypothesis is to directly test this effect in Chinese speaking children. Children in Mainland China begin to receive formal literacy instruction in first grade (Chen, Hao, Geva, Zhu, & Shu, 2009). However, their understandings of meaningful radicals/compounds in written scripts are closely related to the development of their morphological awareness, which usually is not sufficiently developed until Grade 2 or even later for both English and Chinese speaking children. For example, Ku and Anderson (2003) reported that morphological awareness was highly related to vocabulary knowledge in second, fourth, and sixth graders in Taiwan and the United States. Furthermore, McBride-Chang, et al. (2005) found in a comprehensive crosscultural comparative study that morphological awareness was similarly associated with vocabulary across second graders in Beijing, Hong Kong, Korea, and the United States. Another study investigating rapid naming speed and orthographic consistency between Grade Three English-speaking Canadian children, Greek-speaking Cypriot children, and Chinese-speaking Taiwanese children also found that across languages there were no statistically significant differences in the correlations between rapid naming and reading.

However, morphologically complex words consist of inflections, derivatives, and compounds in different languages. Thus which aspect of morphological awareness is associated with literacy and concept development in a language depends on the specific features of the morphological structure of the language (Chen et al., 2009). In the studies of morphological awareness development in Chinese speaking children, many past researchers have focused on compounding morphology as a significant feature for Chinese reading development. Compound awareness is defined as the knowledge about the meaning and structure of compound words as a combination of constituent morphemes (e.g., bike *zi4xing2che*自 行 车 as the self-moving 自 行 vehicle 车)(Chen et al., 2009). In contrast, orthographic awareness generally refers to children's understanding of the conventions used in the writing system of their language ((Treiman & Cassar, 1997; Wang, Cheng, & Chen, 2006), and is also argued to be important for Chinese reading. In a recent study using the compound analogy task (See the *behavioural measures* in Methods), Chen and colleagues found that compound awareness is a significant predictor of Chinese character reading and is significantly associated with vocabulary development among Grade 1 and 3 Chinese monolingual children (Chen et al., 2009). In another study using an orthographic choice task (See also the *behavioural measures* in Methods), Wang and colleagues found that orthographic awareness was the most powerful predictor for Chinese character reading in Grade 2 and 3 Chinese children (Wang, Perfetti, & Liu, 2005).

Examining the impact of linguistic conventions for children who have much more experience with morphological/verbal cues (but still less than adults) than orthographical/written cues in everyday nouns could thus provide us a unique chance to verify previous findings from adults about these two kinds of category labels.

A critical hypothesis in our model proposed with the adults' data is that both the difference between the language labelling conventions and speakers' experience with it influence the comparing and accessing of the linguistic category information during category judgement. We already showed that English speaking adults who have less morphologically transparent labels in the language will show difference typicality effect in the brain compared with Chinese speaking adults, that is, the labelling conventions in the language matters. However, we still did not show how the speakers' experience could influence the effectiveness of linguistic labels, which is the focus of our children study here.

First of all, since Grade 2 or 3 Chinese speaking children have relatively less experience with orthographical cues in the written language, we might expect to find a stronger N300/N400 typicality effect for orthographically transparent items from Chinese speaking children than from Chinese speaking adults, perhaps in parallel to what was found for nontransparent items in English speaking adults. Secondly, since Chinese speaking children at this age already have relatively extensive experience with morphological/verbal language and morphologically transparent labels, we might expect to find a weaker N300/N400 typicality effect for morphologically transparent items than orthographically transparent items. In contrast to Chinese speaking adults, however, I would expect morphologically transparent items to generate relatively larger N300/N400 effects. I expect to find a hierarchical structure in English speaking adults, Chinese speaking adults and Chinese speaking children for the influence of linguistic labels on categorization: The nontransparent items in English speaking adults will show the strongest typicality effect, followed by the orthographically transparent items in Chinese speaking children, then the orthographically transparent items in Chinese speaking adults, then morphologically transparent items in English adults and Chinese speaking children. The morphologically transparent items in Chinese speaking adults will show the weakest typicality effect.

I chose this age group for several reasons: first, 8-9 years old, Grade 2 or 3 children already have enough experiences with object categories and their typicality, as evidenced by the behavioural and ERP results in previous studies (Batty & Taylor, 2002), so can successfully name object pictures and perform the task. Secondly, Grade 2 or 3 children have already learned basic Chinese characters and the radicals so can read our word labels in the experiment. Thirdly, 8-9 year old children are much more cooperative in an ERP study than younger children.

Besides behavioral and ERP measurement, I also conducted three morphological awareness and orthographical awareness tasks: *morphological compounding production task* (P. D. Liu & McBride-Chang, 2010), *orthographical radical choice task* (Wang et al., 2005), and *orthographical semantic category task* (Tong, 2008; Tong & McBride-Chang, 2009), to measure the development of children's compound awareness and orthographic awareness. If the influence of morphologically and orthographically transparent labels in children's categorization depends on children's corresponding morphological awareness development, one might expect correlations between the amplitude of the typicality effect and measurement scores. Children who have higher scores on the morphological compounding awareness task might receive a larger influence from morphologically transparent labels and show weaker behavioural and ERP typicality effects for morphologically transparent items, whereas Children who have higher scores on orthographic awareness tasks may receive larger influence from orthographically transparent labels and show weaker behavioural and ERP typicality effects for orthographically transparent items.

Method

Participants

Twenty-three (12 females, *Mean* age = 8.47 years, *SD*= 0.51) Third Grade native Chinese-speaking children, all right-handed with normal vision, participated in this study and received payment. Two female participants were excluded from further analysis because of uncorrectable noise and eye-blinking during EEG recording.

Stimuli

Twenty-eight greyscale object pictures for 10 categories were selected from Study 2 (C. Liu et al., in press) (Fig.1 C and D). To ensure consistency across categories for children, every child completed a typicality rating survey after the ERP recording by first naming the item 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?). All item pictures received >80%

naming accuracy at the item-level from Chinese children. The typicality rating results for the ten categories in Chinese (Table 5) revealed no significant interactions between typicality and label type, thus ensuring the comparability of the stimuli across age group and conditions in the current study.

Procedure and Task

The procedure, apparatus and task of the study are the same as in our previous Study 1 (Liu et al, in press).

Behavioural Measures

Morphological compounding production task. This task was designed to test how well children could use their knowledge about morphemes and morphological structure to produce novel words that are not used in real speech (P. D. Liu & McBride-Chang, 2010). In this task, children were asked to produce a novel word in response to an aurally presented question. Children were encouraged to produce the novel word that could most properly express the meaning conveyed by the question/scenario. Only when they could retrieve the critical morphemes and combine them according to the specific structure indicated by the sentence could children be considered to have produced the model answer. For example, one question was, “What should we call a monster that eats iron?” The best answer, as determined through pilot testing, was 吃铁怪 (chi1 tie3 guai4, iron-eating monster)(P. D. Liu & McBride-Chang, 2010). All responses were rated on a 0- to 4-point rating scale by two trained raters (interrater reliability Cronbach's $\alpha = 0.96$). A 4-point answer included all critical morphemes and a correct and succinct structure; 3 points were allotted for a response that included unnecessary morphemes, that is, a correct but redundant structure (e.g., 吃铁怪兽where 兽 is redundant, for the above example); 2

points were given for an answer that missed critical morphemes, that is, an incomplete structure (e.g.,铁怪 [tie3 guai4, iron-monster]); 1 point was given for a response with some of the critical morphemes but an incorrect structure (e.g., 怪铁 [guai4 tie3, monster-iron]); 0 points were allotted for an unrelated response or no response (P. D. Liu & McBride-Chang, 2010). The test contains 37 items in total (Appendix 7). The results can be found in Table 6.

Orthographic radical choice task. In this task, children wiere presented with a pair of noncharacter stimuli on a card. They were instructed to choose the one that looked more like a real character (Wang et al., 2005). Forty items were included across two conditions with 20 items each (Appendix 8). The first condition measured children's sensitivity to the legality of the radical position, whereby one of the pairs of stimuli contains a component radical in an illegal position, for example, in the pair  and ,  contains a legal radical in an illegal position. The second condition measured children's sensitivity to the legality of the radical form. One of the pairs of stimuli contained a component radical with an illegal form, for example, in the pair  and, contains an illegal radical. Illegal radicals were created by adding, deleting, or moving a stroke from one location to another within a legal radical. The accuracy of forty items will be measured as the independent variable. The results can be found in Table 6.

Orthographical semantic category task. This task tapped children's sensitivity to cue the meaning of radicals in semantic-phonetic compound characters (Tong, 2008; Tong & McBride-Chang, 2009). It was comprised of 37 items ranked in order of increasing difficulty (Appendix 9). There were 16 target pseudocharacters consisting of bound radicals and another 21 including free radicals. Each item consisted of one target pseudocharacter and four colorful pictures of concrete objects or concepts from which to select the correct answer. The correct answer was the choice of the picture, from among four, that best represented the meaning of the given pseudocharacter. For example, children were presented with a target pseudocharacter, e.g., 鮟, along with four pictures showing a watermelon, a flower, a dragonfly, and a fish, and they were asked to choose the picture that could best represent the meaning of the target pseduocharacter. The correct answer for this example was the picture of the fish because the radical 魚 of the pseudocharacter 鮟 signifies the meaning of fish or a fish-related concept. There was one practice item and testing stopped when children failed five consecutive items. Each correct selection was credited one mark, and the maximum possible score for this task was 37. The results can be found in Table 6.

EEG Recording

Presentation of the stimulus was controlled with E-prime and displayed on a 20 inch monitor from 54cm away. The electroencephalogram (EEG) was recorded from 128 scalp sites (whole skull-referencing) using128 HydroCel GSN net with Amp 3.0 (EGI). The electro-oculogram (EOG) was monitored with six electrodes placed bilaterally in the external canthi (128 and 125), supraorbital (26 and 8), and infraorbital (127 and 126) regions. All interelectrode impedances were maintained below 50 kΩ. The EEG and EOG were amplified using a 0.01- 100 Hz bandpass and continuously sampled at 500 Hz/channel for off-line analysis. Trials with EOG artifacts (mean EOG voltage exceeding ± 200 µV) and those contaminated with artifacts due to amplifier clipping, bursts of electromyographic (EMG) activity, or peak-to-peak deflection exceeding ± 150 µV were excluded from averaging. Data were recorded and processed using Net Station 4.3 (EGI software). After acquisition, the data were lowpass filtered below 20 Hz. The continuous EEG was segmented into an epoch starting at 100 ms before the onset of the stimulus and lasting until 1200 ms after stimulus onset. Segmented files were scanned for artifacts with the Artifact Detection toolbox in NetStation 4.3 using a threshold of 70 μV for excessive muscular activity, eye blinks, and eye movements. Segments containing eye blinks or movements as well as segments with more than 20 bad electrodes were rejected. Within each segment, electrodes with either an average amplitude of greater than 200 μV or difference average amplitude of 100 μV were also discarded from further processing. Finally, particular electrodes were rejected if they contained artifacts of any kind in more than 50% of the segments. Artifact-free segments for correct responses were averaged separately over the 1200-ms epoch across subjects and re-referenced against the average of all electrodes.

Results

Behaviour Results

Trials with a response time > 2500 ms or < 200 ms were deleted as outliers (479 trials, 5.73% of all responses). Table 5 shows the the accuracy, reaction time (RT) and typicality rating data for Morphologically *vs.* Orthographically transparent labels and Atypical *vs.* Typical items in *Yes* responses. A Typicality (Typical *vs.* Atypical) by Label type (Morphological *vs.* Orthographic) ANOVA found that the main effects of Typicality were significant for both RT and typicality rating data, but not for accuracy data. Participants made rated higher, $F(1, 20) = 69.30$, $P < 0.001$ and responded more slowly, $F(1,20)=15.51$, $P < 0.001$ for Atypical than Typical items. They also rated higher, $F(1,20) = 26.53$, $p<0.001$ for morphologically transparent items than orthographically transparent items. The post-hoc analysis with Bonferroni corrections in RT showed significant typicality effect for both morphologically transparent items, $P < 0.05$ and orthographically transparent items, $P < 0.01$. No significant Typicality by Label type interactions was found.

ERP Results

We did not identify clear N300 and N400 ERP components in the grand average wave forms. Instead, the late slow wave (LSW) ERP components (D. Liu, Meltzoff, & Wellman, 2009; D. Liu, Sabbagh, Gehring, & Wellman, 2009) were presented for both the morphologically and orthographically transparent items (Fig.12). Thus, we focused our analysis on the LSW component, which was measured as the mean amplitude in the 600-900 ms interval. All epochs were measured following the onset of the target picture, relative to a 100ms pre-stimulus baseline.

As in the Study 1 (Liu et al, in press), we focused our analysis on the horizontal line encompassing the bilateral frontal region (F7, F5, Fz, F6, and F8, corresponding to 38, 27, 11, 123, 122 in the 128 HydroCel GSN Net, respectively). Voltage data for peak amplitude of the LSW components from two left-right pairs, F5 - F6 and F7 - F8 for the *Yes* responses were used in a Typicality (Typical *vs.* Atypical) by Side (Left [F5 or F7] *vs.* Right [F6 or F8]) by Label type (All *vs.* Morphological *vs.* Orthographic) repeated measures ANOVA with Bonferroni corrections for post-hoc analyses (Fig.12).

The LSW component showed significant or marginal main effects of Side in both the F5 - F6 and the F7 - F8 pairs. The left side (F5 or F7) showed larger LSW than the right side (F6 or F8) (F5 - F6 pair: M = 2.61 and 7.45 μV, respectively, F (1, 20) = 4.02, $P = 0.06$; F7 - F8 pair: M = -1.41 and 7.75 μV, respectively, F (1, 20) = 8.35, P

< 0.01). In addition, the F5-F6 pair also showed a significant main effect of Typicality, such that Atypical items showed larger LSW than Typical items, *M* = 4.47 and 5.59 μV, respectively, *F* (1, 20) = 7.93, *P* < 0.05; whereas the F7-F8 pair also showed a significant main effect of Label type, such that Morphologically transparent items showed a larger LSW than Orthographically transparent items, *M* = 2.78 and3.55 μV, respectively, *F* (1, 20) = 3.60, *P* < 0.05. Most interestingly, both the F5-F6 and F7-F8 pairs showed a significant or marginally significant Typicality by Label type interaction for the LSW component, F5-F6, *F*(2,40) = 3.17, *P* = 0.05; F7-F8, *F*(2,40) = 12.62, *P* < 0.001, such that orthographically transparent items indeed showed left-lateralized (F7, F5) typicality effects for a negative going LSW component, whereas morphologically transparent items showed no significant LSW differences between Typical and Atypical items at the left frontal electrodes but a significant positive going LSW typicality effect at the right frontal electrodes (F8) (Fig.12). The post-hoc analyses for the typicality effect in each electrode can be found in Fig.12.

We also compared the No response (out of category) with the Typical and Atypical items (combined both label types) (Fig. 12 TOP). Only on significant difference between Atypical and No items was found in the F7 electron, such that Atypical response showed larger LSW than No items, *F* (1, 20) = 8.45, *P* < 0.05. However, we should note that the trials numbers for No response is twice as much as the Typical and Atypical items, thus the difference here might partly be contribute to the trials number differences rather than processing differences.

Correlation Results

The descriptive statistics for all three behavioural measures of children's morphological and orthographical awareness (morphological compounding production task; orthographical radical choice task; orthographical semantic category task) are displayed in Table 6. A bivariate correlation analysis with these three measures and the typicality effects (atypical minus typical) of the six categorization variables (typicality ratings of morphologically transparent items; typicality ratings of orthographically transparent items; reaction time of morphologically transparent items; reaction time of orthographically transparent items; ERP amplitude different at F8 for the morphologically transparent items; ERP amplitude different at F7 for the orthographically transparent items revealed six significant or marginal significant correlations. Table 7 shows the zero-order (Pearson) correlations among all nine measures (P. D. Liu & McBride-Chang, 2010).

Most remarkably, for the morphologically transparent items, the ERP typicality effect at F8 was positively correlated with the reaction time for the typicality effect, $r = 0.50$, $p < 0.05$, but negatively correlated with the morphological compounding production task scores, $r = -0.43$, $p = 0.05$, (Fig.13), indicating that the behavioral typicality effect of the morphologically transparent items is associated with the LSW ERP component at the right frontal regions for Chinese children. Moreover, the ERP typicality effect can be predicted by children's morphological awareness, such that children with higher morphological awareness will elicit smaller ERP typicality effects, and vice versa. In contrast, for the orthographically transparent items, the ERP typicality effect at F7 was marginally negatively correlated with the reaction time typicality effect, $r = 0.38$, $p = 0.09$ (Fig.13), indicating that the behavioral typicality effect of the orthographically transparent items is associated with the LSW ERP component at the left frontal regions for Chinese children.

In addition, orthographical radical choice task scores were marginally negatively correlated with the typicality ratings of the orthographically transparent items, $r =$

-0.38, $p = 0.09$. Orthographical semantic category task scores were negatively correlated with the reaction time typicality effect of the orthographically transparent items, $r = -0.60$, $p < 0.01$ (Fig. 13; Table 7), such that children with higher orthographical awareness elicited smaller behavioral typicality effects, and vice versa.

In summary, the correlation analyses indicate that the behavioral typicality effects of morphologically and orthographically transparent items have different brain correlates. Specifically, morphologically transparent items are associated with a positive going LSW component in the right frontal region whereas orthographically transparent items are associated with a negative going LSW component in the left frontal region. Additionally, children's morphological awareness can predict the ERP typicality effect of the morphologically transparent items, while children's orthographical awareness can predict the behavioral typicality effect of the orthographically transparent items.

Discussion

Unlike our previous studies with Chinese speaking adults, the behavioural typicality effect in Chinese speaking children was only found in the reaction time data, which is in line with previous studies in children showing that behavioural measurements are less consistent for children than adults (Mervis & Pani, 1980). Such an attenuated typicality effect was also found in ERP measurement. Compared with adults, children did not elicit clear N300 and N400 ERP components but a bilateral late slow wave (LSW) ERP component, which was also found in previous ERP studies with English speaking children (D. Liu, Meltzoff et al., 2009; D. Liu, Sabbagh et al., 2009).

ERP results showed different patterns of the typicality effect than that of adults' results (see Study 2). As we have hypothesized, when the morphologically and orthographically transparent items were combined together, the LSW showed significant differences between Typical and Atypical items in the left frontal electrodes (F5) (Fig. 10 top), This leads further evidence to the hypotheses that the disappearance of the typicality effect in Chinese speaking adults in Study 1 was due to the influence of linguistic category cues in Chinese. Unlike adults, Chinese speaking children who have much less experience with these linguistic category cues still present the ERP typicality effects in left frontal regions. In addition, we found a significant difference between morphologically transparent items and orthographically transparent items on the LSW component, such that morphologically transparent items showed a significantly positive going LSW typicality effect at the right frontal electrodes (F8) (Fig.10 bottom), whereas orthographically transparent items showed a significantly negative going LSW typicality effects at the left frontal electordes (F7, F5) (Fig.10 top).

Most importantly, we found that the behavioral typicality effects of morphologically transparent items and orthographically transparent items have different brain correlates. Specifically, morphologically transparent items are associated with a positive going LSW component in the right frontal region whereas orthographically transparent items are associated with a negative going LSW component in the left frontal region. Additionally, children's morphological awareness can predict the ERP typicality effect of the morphologically transparent items, while children's orthographical awareness can predict the behavioral typicality effect of the orthographically transparent items.

Interestingly, the pattern of morphologically transparent items in Chinese speaking children is quite similar to English speaking adults, such that the typicality effect of ERP components for morphologically transparent items can be found only in the right frontal region rather than left frontal region.

In our results, we found a left-frontal negative LSW for orthographically transparent items and a right-frontal positive LSW for morphologically transparent items in Chinese speaking children. How could we interpret this interaction between the type of LSWs and the type of linguistic labels? In previous ERP studies, the tendency of LSW has been found to be related to the type of operation performed in working memory. Specifically, conceptual operations elicited the negative going LSW, whereas perceptual operations elicited positive going LSW (Ruchkin et al, 1988, 1990, 1992). For example, studies have found the negative LSW shifts in a verbal concept formation task in which participants were asked to transform letters into Morse codes (Lang et al., 1987).

Thus, the left-frontal negative LSW we found for the orthographically transparent items might indicate some sort of conceptual operations in working memory (e.g., decoding the semantic radical in the Chinese characters and extracting the category information from the radicals. This decoding effect was found only in Chinese-speaking children but not adults because adults have much more experiences on orthographically transparent items that Grade 3 children. In contrast, the right-frontal positive LSW we found for the morphologically transparent items might indicate some sort of perceptual operations in working memory (e.g., identifying and comparing the pronunciation in the morphologically transparent items), which are identical to the preceding category names. There is also another possibility that such a right-frontal positive LSW is an variation of the right frontal N300 and N400 component we found in English-speaking adults, which could reflect additional attention and executive processing (Corbetta & Shulman, 2002; Han et al., 2004) being involved for the analysis of the morphologically transparent category information.

In Study 2, when pictures of items that have morphologically transparent cues to the category were provided, even English speaking adults appeared to make use of these cues to facilitate semantic access and category judgments, as evidenced by the overall decreased RTs for morphologically transparent items and the additional executive attention load (Corbetta & Shulman, 2002; Han et al., 2004) reflected by the relatively greater right frontal activation for morphologically transparent items. I propose that it is because most English nouns are nontransparent rather than linguistically transparent like Chinese nouns (Tardif, 2006; X. L. Zhou, Marslen-Wilson, Taft, & Shu, 1999; Y. G. Zhou, 1978), thus providing pictures of morphologically transparent items is not enough to allow English speakers to circumvent the use of typicality as an aid to categorization.

This hypothesis was directly supported by the similar right frontal ERP typicality effect found in Chinese speaking children. Although Chinese speaking children have already had some understanding about morphologically transparent cues, their overall learning experiences are much less than Chinese speaking adults but similar to English speaking adults, who have less morphologically transparent examples but more overall learning experiences due to their age. Thus it is not surprising that they elicit similar patterns of typicality effect with morphologically transparent items in the brain.

These results suggest that 1) In general, with much less influence from their native Chinese language than adults, the ERP pattern showing a typicality effect in Chinese-speaking children is more like that of English adults, such that atypical items elicited large ERP typicality effect in the left frontal regions for nontransparent/orthographically transparent items and right frontal regions for morphologically transparent items; 2) The behavioural and ERP typicality effects in Chinese speaking children are correlated with their performance of their morphological awareness and orthographic awareness tasks. Children with higher scores in the morphological awareness task showed weaker behavioural and ERP typicality effects for morphologically transparent items, thus evidencing their use of this cue in category judgement. Similarly, children with higher scores in orthographic awareness tasks should be better able to access and use the orthographic cues in orthographically transparent labels and thus showed weaker behavioural and ERP typicality effects for orthographically transparent items.

These results also extend the integrated model we proposed in the Study 3 for the data collected from English and Chinese adults. In the model we proposed that the bypass processing for morphologically transparent items in Chinese adults happen in the stage of working memory and long term memory. In addition, we suggested that the efficiency of this bypass processing is influenced by the labelling convention of the language and speakers' experience. The bypass processing of morphologically transparent items is most efficient for Chinese adults because this kind of linguistic label is prominent in Chinese (Tardif, 2006; X. L. Zhou, Marslen-Wilson, Taft, & Shu, 1999; Y. G. Zhou, 1978) and Chinese adults already have enough experience using this type of label in years of practicing, resulting in the disappearance of the typicality effect in both N300 and N400 components and frontal activation. In contrast, the same bypassing of processing is less efficient for English adults because they have much less labels as exemplars in the language, resulting in an attenuated typicality effect in the brain. In the current results of Chinese children, we demonstrated another influence on the efficiencies of bypassing processing, that is, the experience of language speakers, such that Chinese children who have many fewer experience using the morphologically transparent items also showed an attenuated typicality effect in the brain compared with Chinese speaking adults. Thus both the properties of the language and the speaker's experience on it play a important role on the effectiveness of the influence of linguistic label on categorization processing.

Chapter VII

Conclusion

Our ERP and fMRI data suggest that language has a direct impact on categorization processes. Speakers of different languages show different patterns of reliance on typicality during category judgment tasks. Moreover, within a language such as Chinese, objects that have category labels embedded in their names also show different reliance on typicality during these same types of category judgment tasks. These results are impressive not only because they show the impact of language on categorization both between and within languages, but also because we provided pictures of the objects and not their labels. Thus, the effects of language on categorization hold even with pictorial stimuli. In addition, the absence of a typicality effect in the N300/N400 ERP components and frontal activation have not been reported in previous studies and is thus a unique and intriguing aspect of the present set of studies.

In the present studies, the category judgment 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 stimulus shown before it, we assume that all participants, both in China and in the US, engaged in some sort of semantic access before the final decision was made. Of great interest, however, is whether participants in the two cultures engaged in the same type or level of semantic access and/or whether they engaged in additional semantic processing of both the pictorial stimulus and the word label (e.g., car *jiao4che1*) given to the stimulus. 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 processing, evidenced by the presence of a typicality effect at both the N300/N400 components and left frontal activation for both morphologically transparent (attenuated) and nontransparent items, in order to facilitate judgments in this task. In contrast, speakers of Chinese were able to bypass this additional semantic processing because of the presence of the category name in the common morphological (e.g., car *jiao4che1*轿 车) label that speakers accessed even when shown a relevant picture. However, they were not able to fully bypass this when given pictures with orthographically transparent information which, as discussed above, appears not to be available for use with much later in development and does not share the additional phonological cues provided by morphologically transparent nouns.

Our findings have enriched our understanding of categorization and will allow us to develop more parsimonious theories to explain this fundamental human ability from neurological, cross-cultural and developmental perspectives. Although traditional behavioural 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 in the brain, probably because of the lack of sensitivity of those behavioural methods, the PDP approach (T. T. 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 modelling, 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 and other types of cues. Both our ERP and fMRI data support the critical role of training and experience on forming representations in semantic memory. In our ERP data, we found that English adults and Chinese children actually demonstrate similar typicality effects when measured using ERP methods. Although English adults have more experience producing and accessing categorical information than Chinese children, English nouns are less likely to contain pronounceable and morphologically transparent cues than Chinese nouns. Thus, one way of thinking about these data is to consider that English-speaking college students may have had roughly comparable amounts of training using morphological cues to categories as Chinese-speaking third graders, although the consistency of these cues is still higher for Chinese-speaking children than for English speaking adults. Similarly, our fMRI data show a pattern such that the morphologically transparent items for adult English speakers and orthographically transparent items for adult Chinese speakers actually activate similar brain areas in the inferior and medial frontal gyrus.

Our results also support the proposed multi-level of categorization systems in the brain (E. E. Smith & Grossman, 2008). In our fMRI study across languages, the reduction of frontal activity for English morphologically transparent items, relative to nontransparent items, and the increase in frontal activation for Chinese orthographically transparent items, relative to morphologically transparent items, together demonstrate the influence of multiple levels of linguistic cues in categorization that are dependent on the prevalence and explicitness of such cues in the language. These results thus are in line with previous findings about different categorization systems in the brain (Grossman et al., 2003; Grossman et al., 2006; Grossman et al., 2002; Koenig et al., 2002; Koenig et al., 2007; Koenig et al., 2008).

Given the present findings together with previous findings, I propose that 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 information, 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. 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.

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 also 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.

To return to Whorf, and perhaps to Shakespeare – "What's in a name? That which we call a rose by any other name [may] smell as sweet." However, our data show that when a rose, or a canola, is called by a name that includes category information (e.g., *you2cai4hua1* 油菜花, or "canola flower"), it changes the way we think and the ways our brains access semantic information. These data also demonstrate when these differences occur in the brain. They suggest that typicality is a useful heuristic for deciding whether a rose is a flower when one's language does not regularly embed category-level terms in the labels for members of the category. Both the patterns of similarity and the patterns of divergence in the N300 /N400 ERP components and frontal activation between English and Chinese speakers and between the two different types of nouns for English and Chinese speakers in this categorization task suggest that category level judgments can undergo differences in processing at early- to mid-stages of stimulus processing (approximately 300 to 400 ms after stimulus presentation at the left inferior frontal gyrus), and yet show similarities at later stages of processing (e.g., LPC after 500 ms at the medial/superior frontal gyrus) and in behavioral responses. In other words, different brain processes can produce similar behavioral outcomes. At the least, these data demonstrate that whether one finds support for the linguistic relativity hypothesis (and for the typicality effect) may depend on the strength and pervasiveness of the linguistic information provided.

An interesting implication of these findings pertains to patients who present with symptoms of semantic dementia (SD) (Basso, Capitani, & Laiacona, 1988). In studies of English-speaking individuals, patients diagnosed with SD tended to not only have general word-finding and other semantic and conceptual difficulties, but a specific regularity in their behavior involving an over-reliance on information that is “typical” of the category or knowledge base being tested. This is true not only for words, but for pictures of real and imaginary animals as well as for real and nonsense words with typical and atypical spelling patterns (Hauk et al., 2006; T. T. Rogers et al., 2004; Woollams, Cooper-Pye, Hodges, & Patterson, 2008).

A particularly interesting characteristic of SD patients is that they usually show a conjunction of semantic and lexical deficits. The most pervasive and self-evident impairment observed in SD is a marked anomia in semantic tasks, specially object naming (T. T. Rogers et al., 2004). First, patients often produce a name that is correct but is more general than the label usually given by age-matched controls for the same object (e.g., animal instead of dog) (Hodges et al.,1995; Warrington, 1975). Second, highly familiar or typical names are often inappropriately extended to semantically related objects (e.g., dog for pig, goat, and sheep)(Hodges et al., 1995). These patterns have been well documented in longitudinal case studies and cross-sectional group studies (Lambon Ralph, Graham,Ellis, & Hodges, 1998; Lambon Ralph et al., 2001).

Similarly, SD patients also have difficulty in lexical decision tasks with word stimuli that parallel the deficits they show in semantic naming tasks with pictorial stimuli (McClelland et al., 2009). As their semantic disorder progresses, they make progressively more errors on words that are of low typicality, especially when they are also of low frequency (Knibb & Hodges, 2005; Woollams et al., 2008). A similar typicality effect is also seen in lexical decision. When asked which of the two letter strings seize and seese is a real word, severe SD patients actually prefer the incorrect but more typical spelling seese (Knibb & Hodges, 2005). This effect is similar to an effect seen in an object-decision task (Rogers, Lambon Ralph, Hodges, 2004), in which participants were asked to choose between a real elephant with large floppy ears of the kind that one only sees on elephants and an otherwise identical elephant with smaller, more typical ears taken from a monkey. In this task, as with the lexical decision task, severe SD patients tended to choose the pseudo-elephant over the real one.

The similarity between these semantic and lexical deficits in SD patients thus provides us a unique chance to investigate the relationship between language and thought and suggests further that there is a similar set of processing structures that underlies semantic processing of objects and words (McClelland et al., 2009).

However, most data on SD collected thus far are from English speakers. Given all these language differences between English and Chinese, as well as all the behavioural and neurological findings I presented here, a very promising question then is whether Chinese patients with SD would also show patterns in their symptoms and behaviours that are similar to their English-speaking counterparts, or whether the morphological and orthographical information provided in Chinese nouns could be used to help ameliorate these symptoms. Answers to this question will help us develop a single-system universal approach to semantic and lexical processing that could account for both the cross-cultural neurophysiological, behavioural and developmental data in the present study and the data that has emerged and can be collected for SD patients from English- *and* Chinese-speaking background. Data from SD patients could also help us clarify the processing stages we proposed in our model.

It is still unclear where and when those behaviour impairments in SD patients happened in the brain when performing a category verification task. If we can identify the brain areas and its time course associated with the behaviour impairment in the SD patients' brain, we can then test our model with these data and modify it accordingly.

In addition, it would be helpful to extend these cross-cultural findings to linguistic categories other than object nouns, such as abstract nouns, verbs, or adjectives. Complementary to the differences in nouns, English and Chinese also differ dramatically in the amount of information they provide in action verbs. In English, action verbs (e.g., run, jump) often provide no linguistic indication of which object (e.g., foot) that these actions are associated with; whereas in Chinese, the object name (e.g., foot/*zhu2*足) is often embedded in the action verb as an orthographic radical (e.g., run/*pao3*跑, jump/*tiao4*跳), thus providing explicit information to help Chinese speakers learn the association between objects and actions. Studies have shown that English and Chinese infants and preschoolers differ greatly in verb learning (Tardif, 1996, 2008). I plan to conduct a behavioural/neuroimaging study using a word-picture association paradigm (e.g. first show the participant a verb RIDE/*qi2*骑 , followed by a picture of a man riding a horse/*ma3*马 or riding a bicycle/*zi4xing2che2*自行车) to test the behavioural and neural influences of verbs on learning action-object associations in English and Chinese speakers. I predict that Chinese speakers will respond differently for action-object pictures (e.g., horse/*ma3*马) that are consistent with the linguistic action-object cue provided by the verb (e.g., ride/*qi2*骑) than those that are not (e.g., bike/*zi4xing2che2*自行车). In contrast, English speakers will not show such a difference.

In sum, we believe it is fruitful to investigate approaches to cognition, and categorization in particular, that invoke a universal mechanism integrated with psycholinguistic and other experiential factors as the foundation for knowledge storage and generalization. The current thesis has had some success in linking research from these different methods, using cross-cultural comparisons, neurological evidence, and a developmental perspective, under a common theoretical framework based on the principles of parallel distributed processing and multiple processes involved in categorization. This dissertation clearly provides more questions than it answers, but the data it does provide guides us in framing several new questions as we move forward to an improved understanding of how activation of semantic and other forms of knowledge dependent on categorization may be influenced by language.

Figure 1

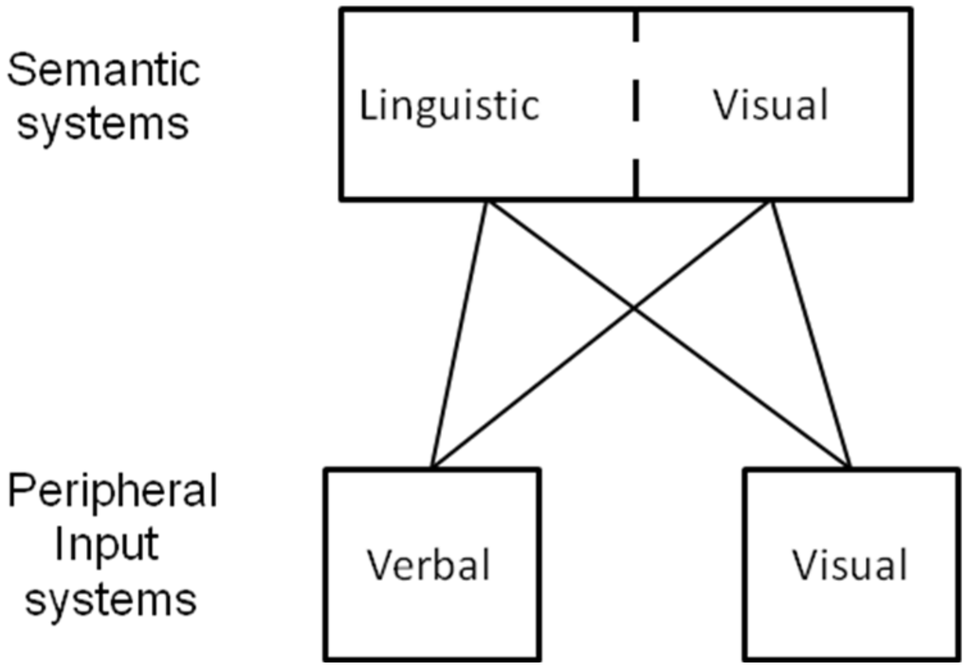


Figure 1. *Schematic diagram of the parallel distributed processing model of semantic memory* (Farah and McClelland, 1991).

Figure 2

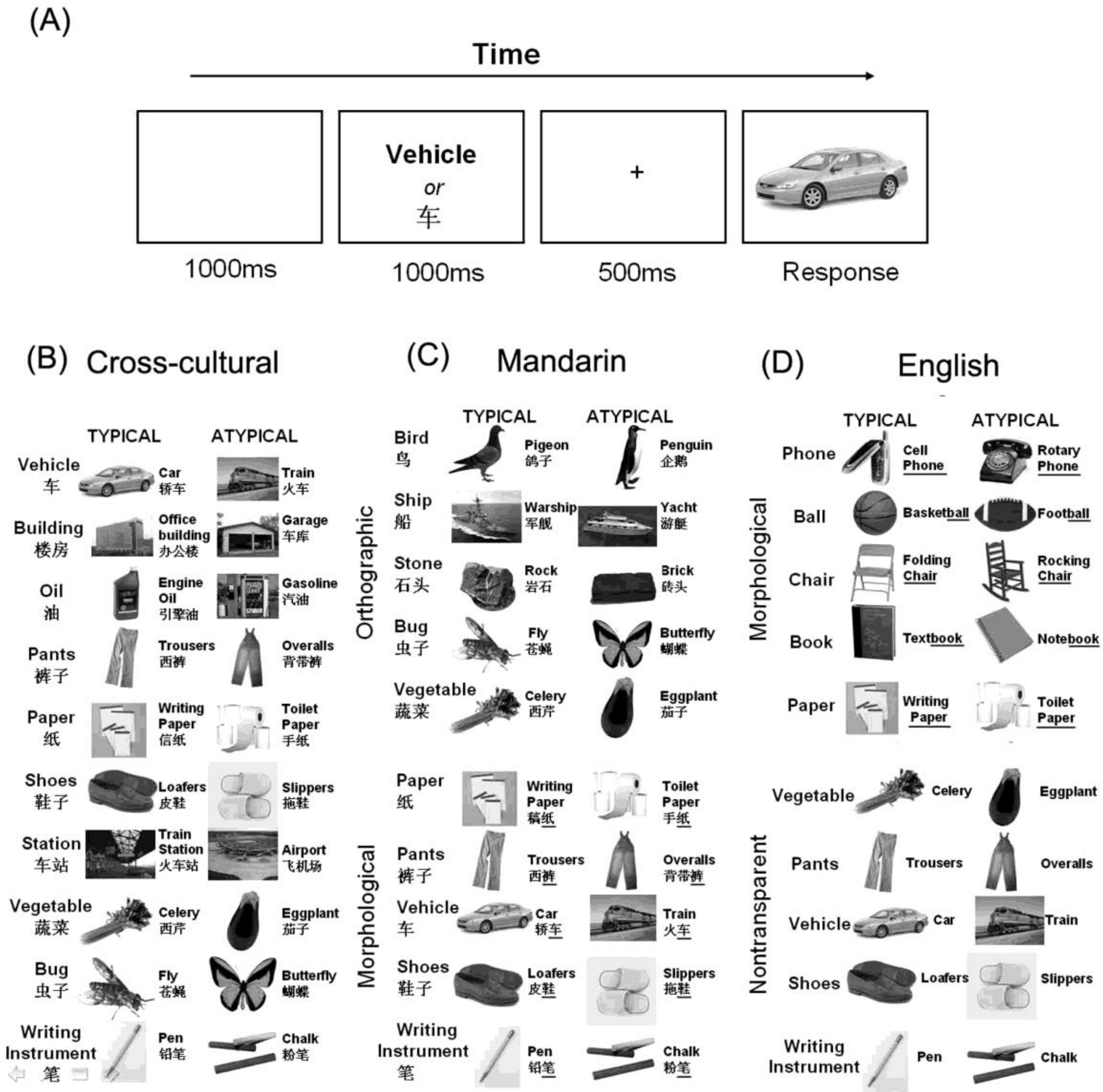


Figure 2. ***(A)*** *Experimental procedure and* ***(B, C, D)*** *Materials with labels and grayscale photographs in the three studies.*

Figure 3

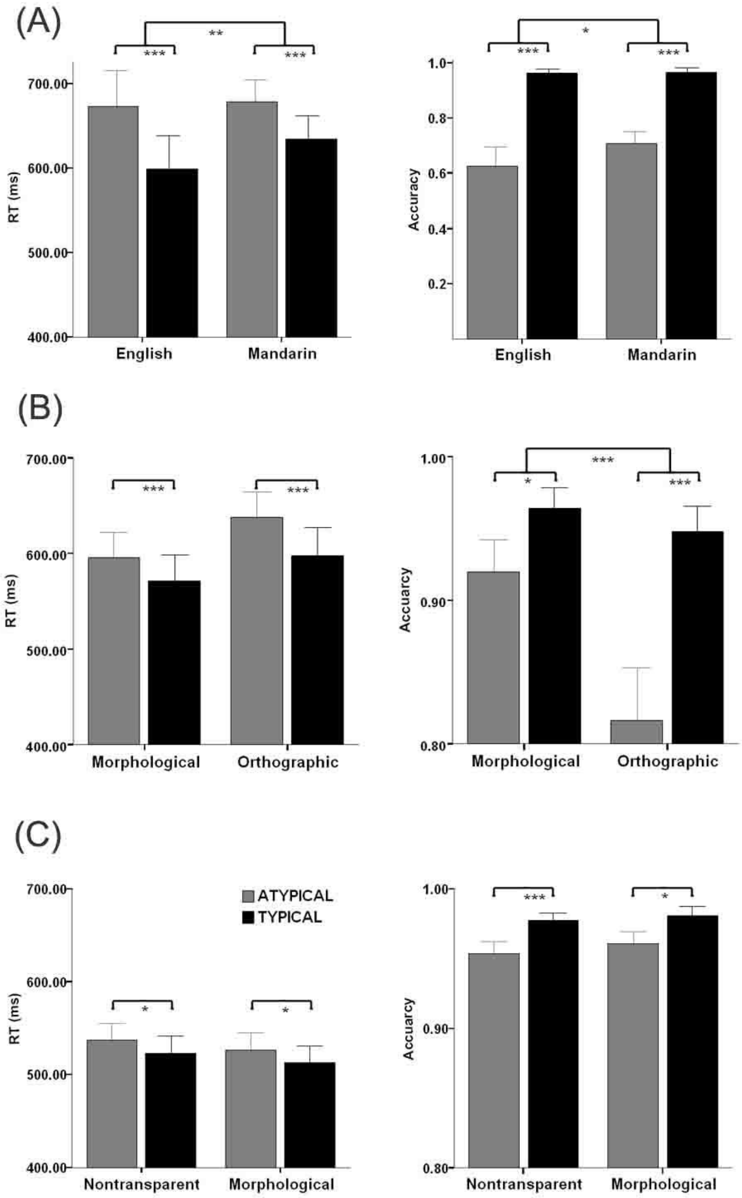


Figure 3. ***(A)*** *English- and Chinese-speaking participants' mean accuracy and RT data (error bars show 2 SE) in Study 1 for typical and atypical pictures in response to category-level labels. Reaction time data is presented only for correct responses.* ***(B)*** *Chinese participants' mean accuracy and RT in Study 2 for typical and atypical pictures with orthographically and morphologically transparent items.* ***(C)*** *English participants' mean accuracy and RT in Study 3 for typical and atypical pictures with nontransparent and morphologically transparent items.*

* $P < 0.05$ ** $P < 0.01$ *** $P < 0.001$ for Bonferroni corrected post-hoc comparisons of conditions.

Figure 4

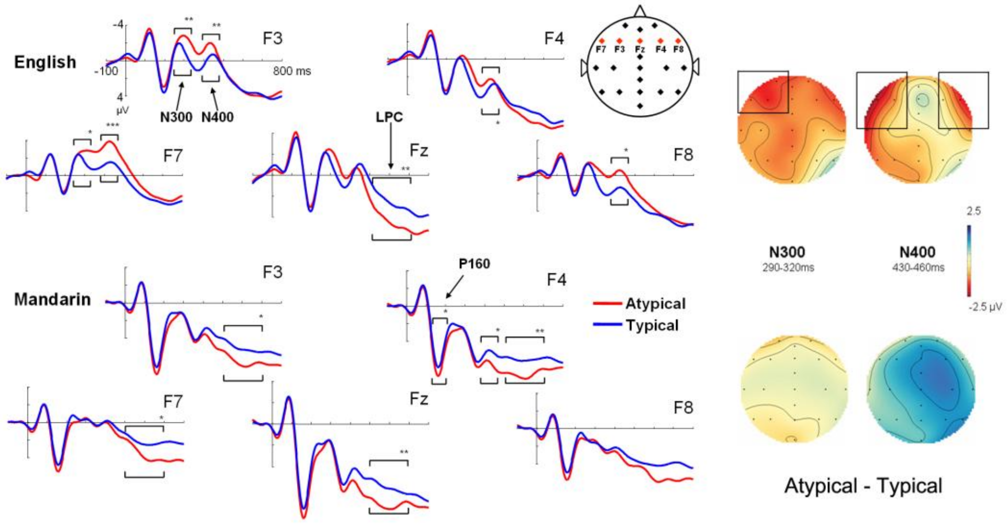


Figure 4. *ERP waves and scalp topographies show divergent difference waves (Atypical-Typical, Correct Yes responses only) for pictures of items as judged by* ***(top)*** *English-speaking participants and* ***(bottom)*** *Chinese-speaking participants.* English-speaking participants show a strong typicality effect in the left frontal region (F3, F7) for the N300 component and bilateral frontal regions (F3, F7, F4, F8) for the N400 component. In contrast, the expected N300 and N400 differences were not found for Chinese-speaking participants. Both English and Chinese speaking-participants also show a strong typicality effect in the middle frontal region (Fz) for the LPC component, for which Chinese speaking participants showed a more widespread distribution in the bilateral frontal regions (F3, F4, F7).
* $P < 0.05$ ** $P < 0.01$ *** $P < 0.001$ for Bonferroni corrected post-hoc comparisons

Figure 5

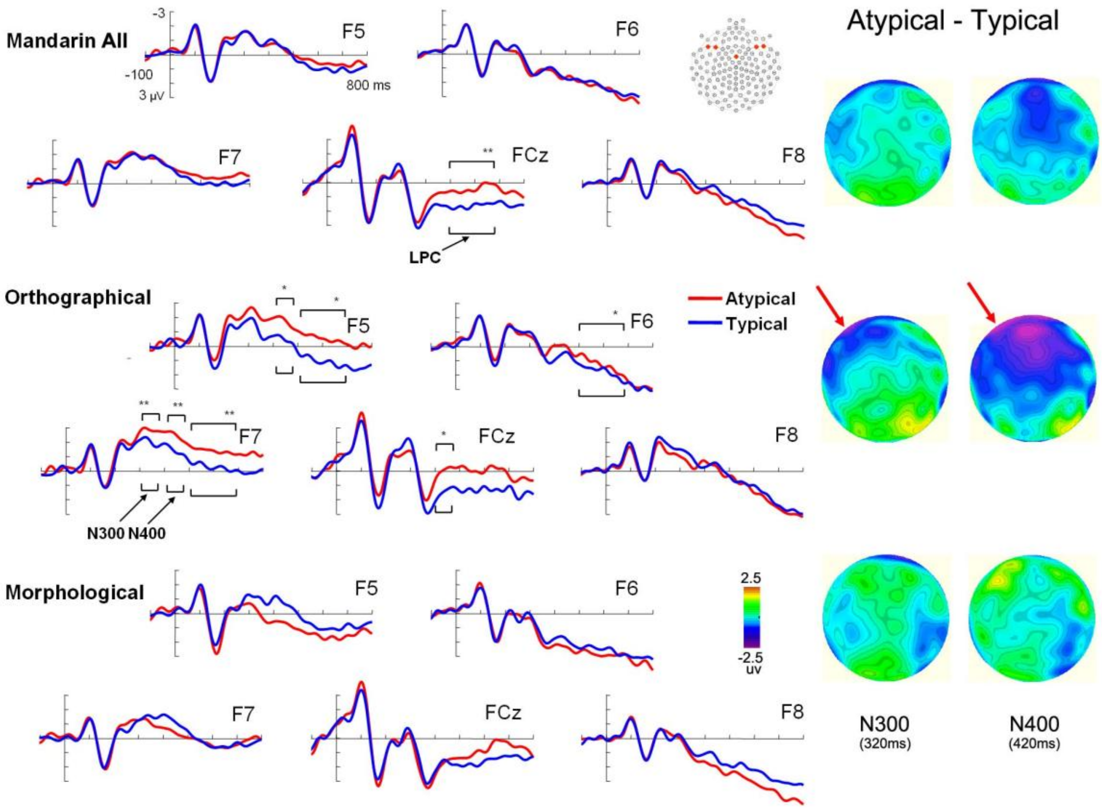


Figure 5. *ERP waves and scalp topographies show divergent difference waves (Atypical-Typical, Correct Yes responses only) for pictures of* ***(top)*** *both orthographically and morphologically transparent,* ***(middle)*** *orthographically transparent only and* ***(bottom)*** *morphologically transparent only typical vs. atypical items after viewing category-level labels.* Chinese-speaking participants showed strong typicality effect in the left frontal electrons (F5, F7) when viewing orthographically transparent items, but no differences when viewing morphologically transparent items for both N300 and N400 components. No significant typicality effect was found when combine them together.

* $P < 0.05$ ** $P < 0.01$ *** $P < 0.001$ for Bonferroni corrected post-hoc comparisons.

Figure 6

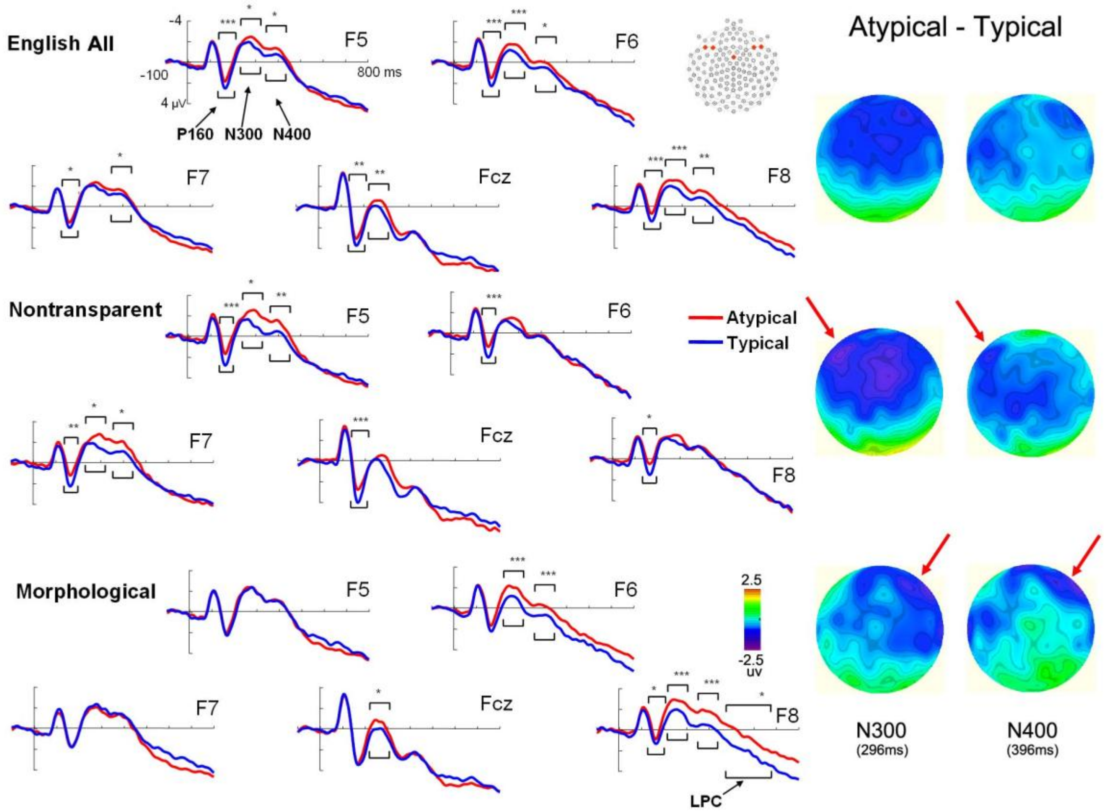


Figure 6. *ERP waves and scalp topographies show divergent difference waves (Atypical-Typical, Correct Yes responses only) for pictures of **(top)** both nontransparent and morphologically transparent, **(middle)** nontransparent only and **(bottom)** morphologically transparent only typical vs. atypical items after viewing category-level labels.* English-speaking participants showed strong typicality effect in the left frontal electrons (F5, F7) when viewing nontransparent items, but in the right frontal electrons (F6, F8) activity when viewing morphologically transparent items for both N300 and N400 components. Significant typicality effect in bilateral frontal regions was found when combine them together.

* $P < 0.05$ ** $P < 0.01$ *** $P < 0.001$ for Bonferroni corrected post-hoc comparisons.

Figure 7

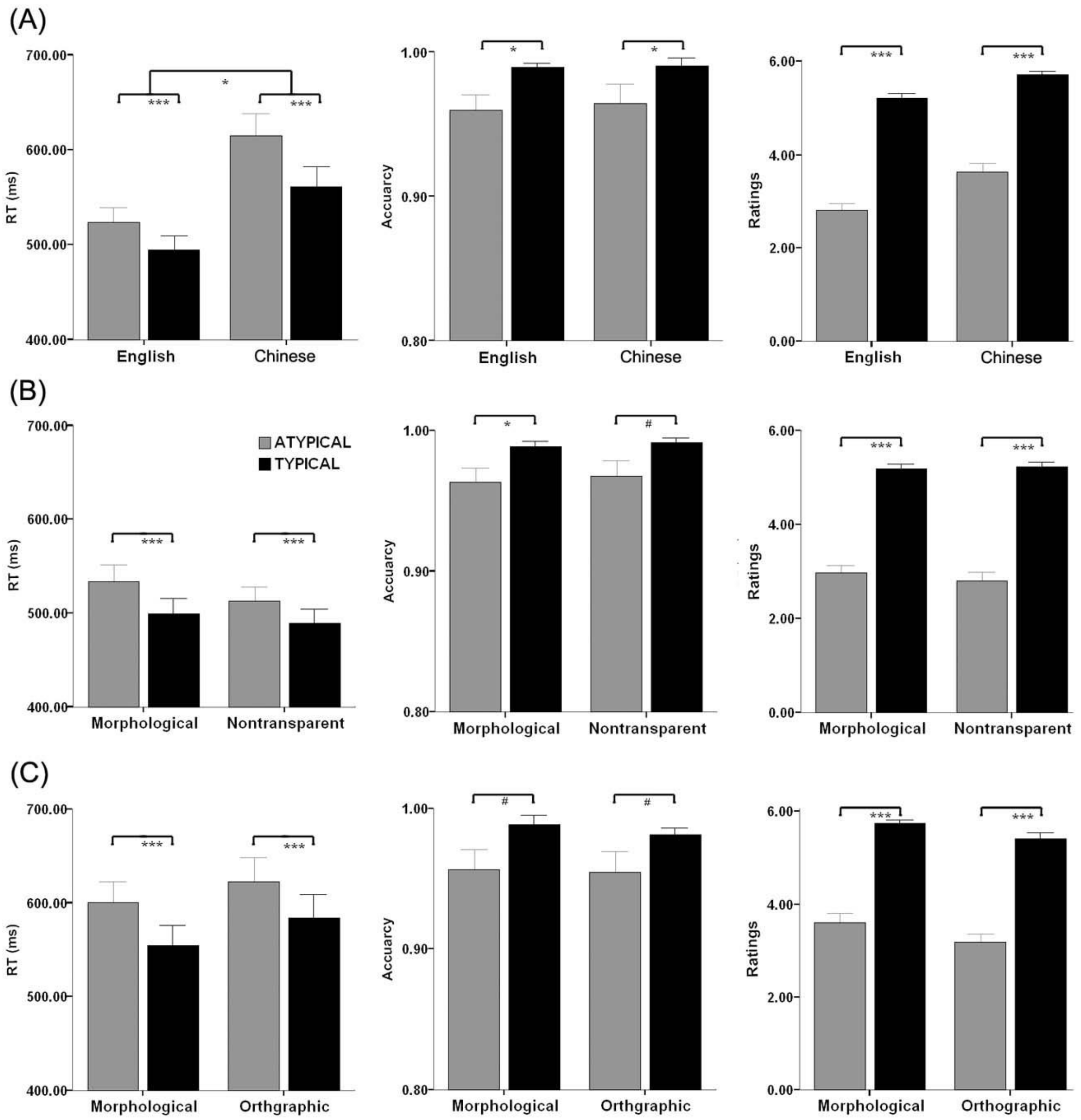


Figure 7. *Behavioral and typicality rating (1-6) results show cross-linguistic differences in the six identical categories shown for English and Chinese speakers* ***(A)*** *, between morphologically transparent and nontransparent items for English speakers* ***(B)*** *and between morphologically transparent and orthographically transparent items for Chinese speakers* ***(C)****. Reaction time data is presented only for correct responses.*

$^{\#} P < 0.01$ $\quad$ $* P < 0.05$ $\quad$ $** P < 0.01$ $\quad$ $*** P < 0.001$

Figure 8

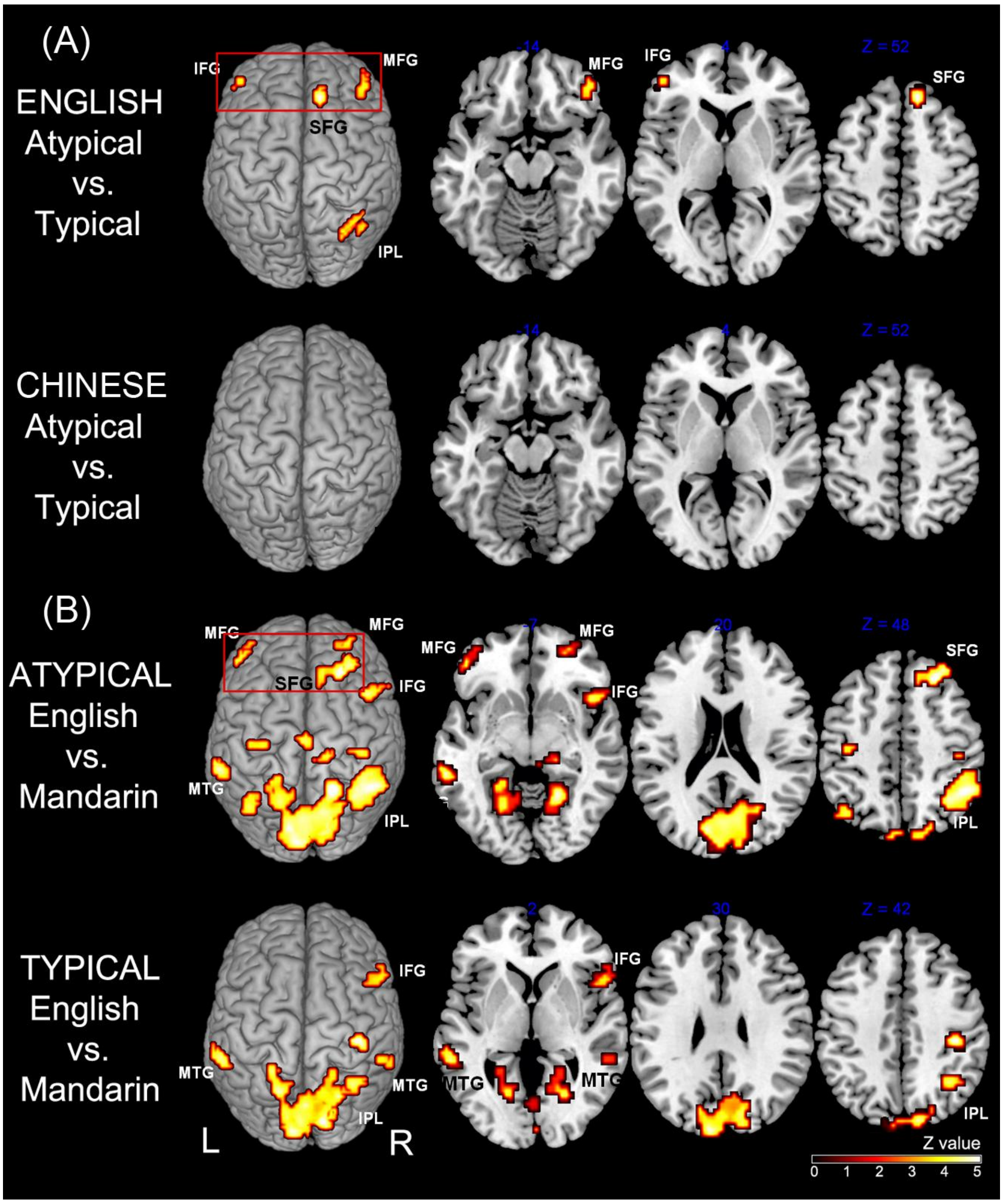


Figure 8. *Brain regions showing activation (atypical-typical) for pictures of items in six identical categories judged by English- and Chinese-speaking participants (Yes responses only). Two sample t-tests, uncorrected voxelwise p < 0.001, corrected p < 0.05 by False Discovery Rate (FDR) method (k > 10 voxels) with small volume correction (SVC) in the regions of frontal and parietal cortex , slices begin with the overall axial view with infinite search depth* ***(A)****. Brain regions showed significant differences in activation between the English and Chinese speakers for both atypical and typical items. (Two sample t-test. Uncorrected voxelwise threshold of P < 0.001, P < 0.05 with FDR corrected. (K > 10 voxels)* ***(B)****.* The typicality effect was associated with the left IFG (BA 46), the right MFG (BA11) and the right SFG (BA8) only in English speakers. Further cross-linguistic contrasts revealed that this frontal activation was associated only with atypical items for the English speakers (Table 1, 2).

Figure 9

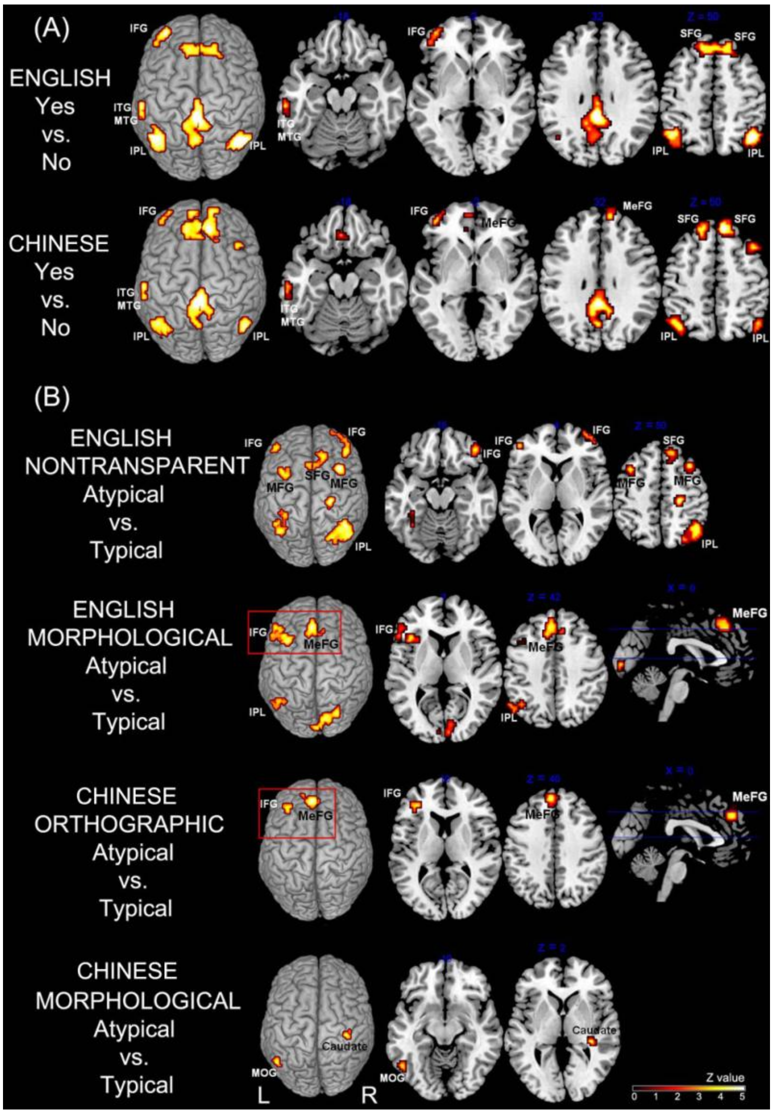


Figure 9. *Similarities in brain activation for pictures of items in all categories that received Yes responses (in category) vs. No responses (out of category) from English- and Chinese-speaking participants. Two sample t-test, uncorrected voxelwise $p < 0.001$, FDR corrected $p < 0.05$ ($K > 10$ voxels)* ***(A)****. Brain activation (atypical-typical) for contrasts among different labels in English or Chinese (Yes responses only). One sample t-test, uncorrected voxelwise $p < 0.005$, FDR corrected $p < 0.05$ ($k > 20$ voxels) with SVC (Fig.3) (FDR corrected $p < 0.1$ for the contrast of Chinese morphological (Atypical vs. typical)* ***(B)***. English and Chinese speakers showed different patterns of activation for the typicality effect dependent on the type of linguistic label (Table 4) despite almost identical patterns for the Yes vs. No contrast (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).

Figure 10

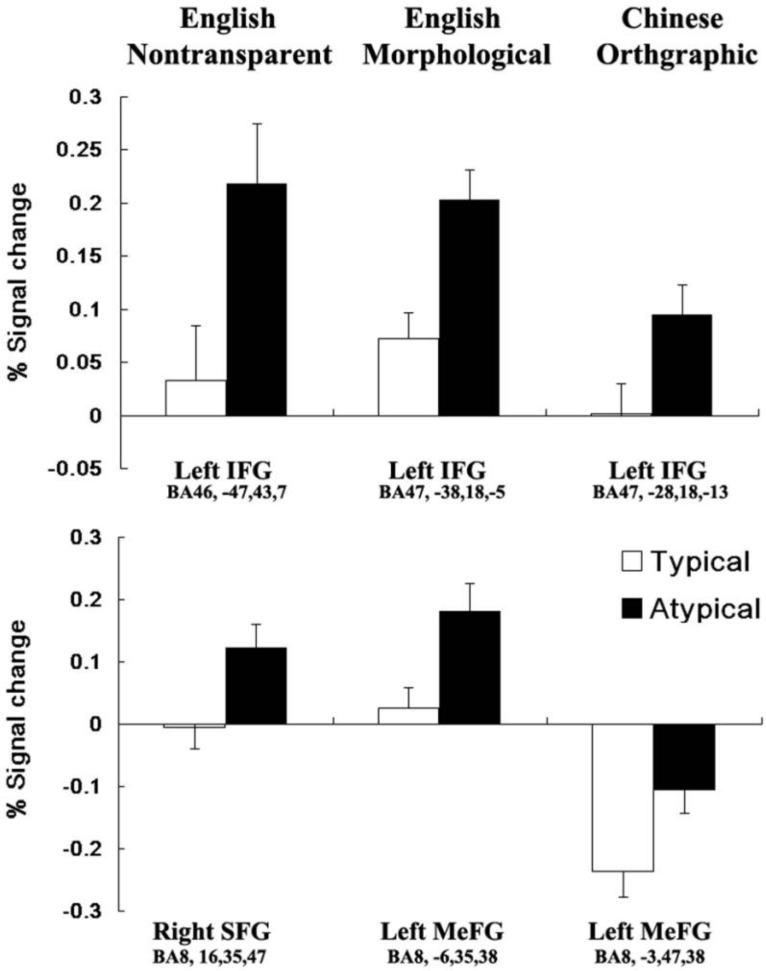


Figure 10. *% Signal change for activation of the atypical vs. typical contrast in the left IFG and SFG/MeFG clusters for English Nontransparent items, English Morphologically transparent items and Chinese orthographically transparent items.*

Figure 11

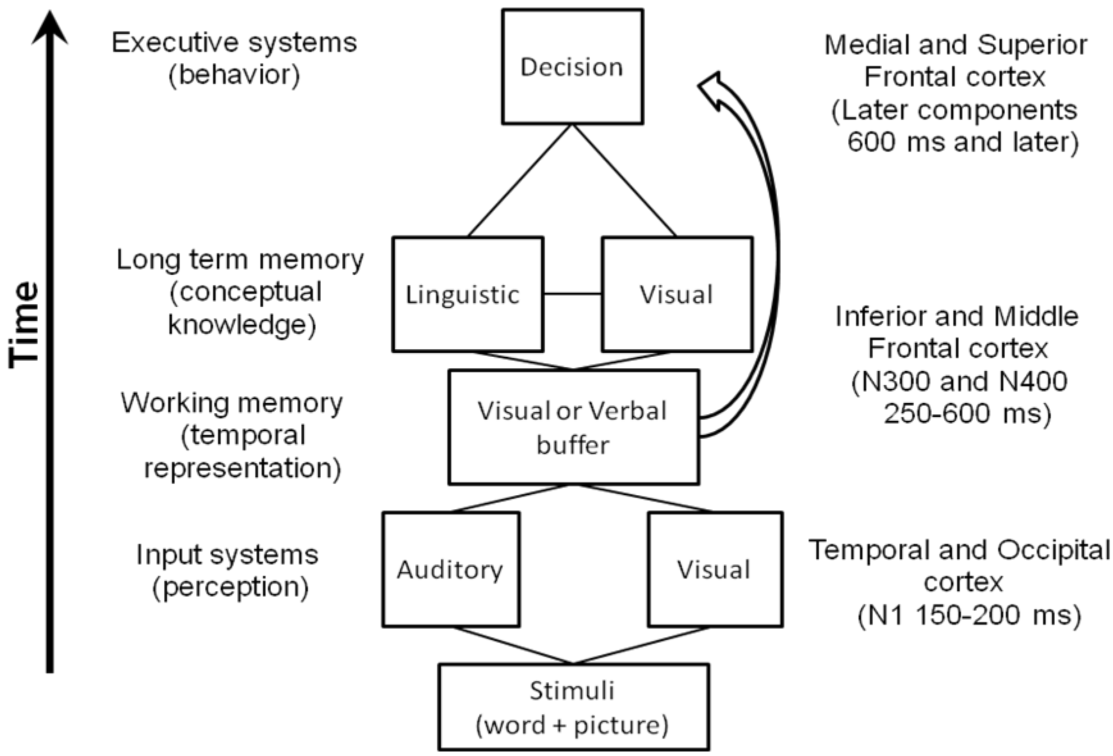


Figure 11. *Schematic diagram of the hypothesised parallel distributed processing model of semantic memory in the brain.*

Figure 12

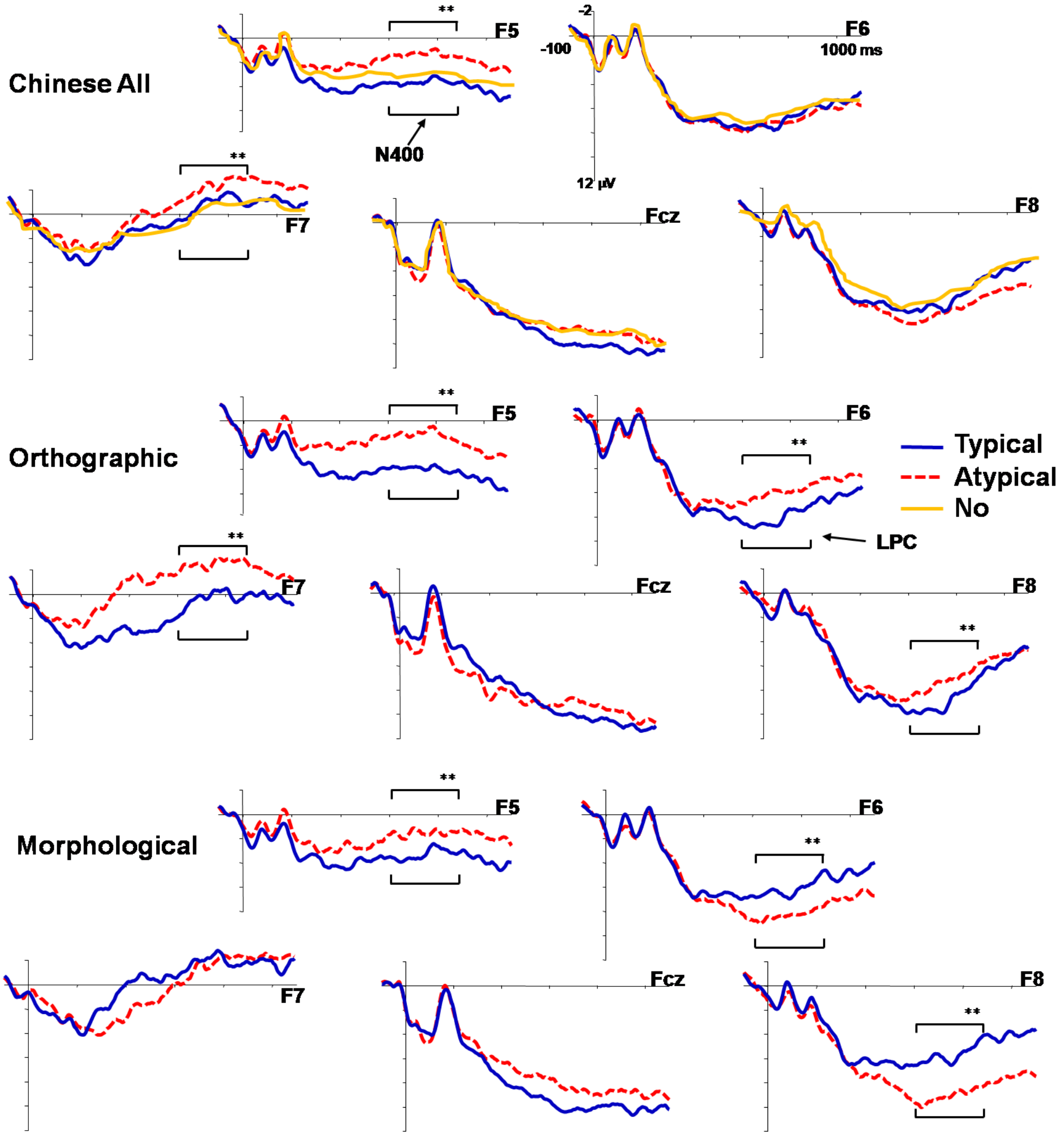


Figure 12. *ERP waves for pictures of* ***(top)*** *both orthographically and morphologically transparent,* ***(middle)*** *orthographically transparent only and* ***(bottom)*** *morphologically transparent only No vs. typical vs. atypical items after viewing category-level labels.* Chinese-speaking childrens showed strong typicality effect in the left frontal electrons (F5, F7) when viewing orthographically transparent items, but in the right frontal electrons (F8) activity when viewing morphologically transparent items for the LSW components. Significant typicality effect in the left frontal regions was found when combine them together.

[#] *P* < 0.1 * *P* < 0.05 ** *P* < 0.01 *** *P* < 0.001 for Bonferroni corrected post-hoc comparisons.

Figure 13

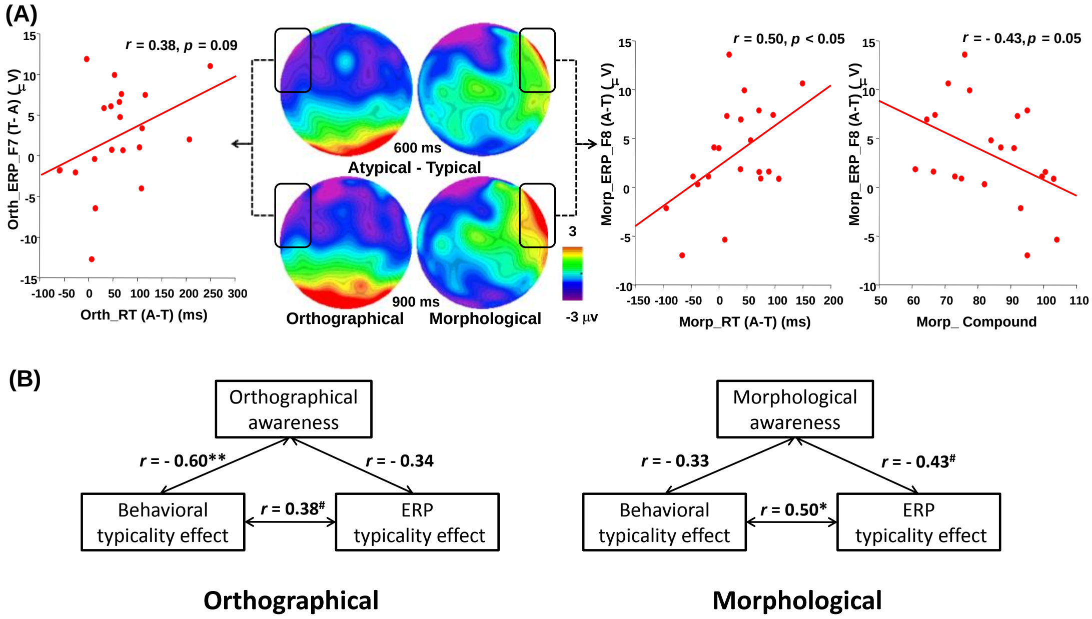


Figure 13. *The voltage maps of the difference wave (Atypical minus Typical) during the time range of 600-900 ms for morphologically and orthographically transparent items **(A)**. The correlations among the morphological and orthographical awareness measurement scores and the behavioural and ERP typicality effects **(B)***. The results showed strong activity in the right frontal regions for the morphologically transparent items but left frontal regions for the orthographically transparent items. For the morphologically transparent items, the ERP typicality effect at F8 was positively correlated with the reaction time typicality effect, but negatively correlated with the morphological compounding production task scores. In contrast, for the orthographically transparent items, the ERP typicality effect at F7 was negatively correlated with the reaction time typicality effect.

$^{\#}$ $P < 0.01$ * $P < 0.05$ ** $P < 0.01$

Tables

Table 1 *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> English 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
English Typical> English Atypical							
L Inferior Occipital G	18	0.01	67	-25	-92	-7	4.36
R Middle Occipital G	19	<0.01	115	28	-93	18	5.14
Chinese Atypical> Chinese Typical							
Chinese Typical> Chinese Atypical							
L Middle Occipital G	18	<0.01	245	-28	-94	1	5.25
R Middle Occipital G	18	<0.01	179	31	-94	1	5.82

Table 2 *Brain regions showing significant activation between English and Chinese speakers for atypical and typical items. (Two sample t-test, uncorrected voxelwise threshold of P < 0.001, P < 0.05 with False Discovery Rate (FDR) corrected for multiple comparisons (K > 10 voxels)).*

Contrast	BA	*P(FDR)*	voxel	*x*	*y*	*z*	*Z*
English Atypical> Chinese Atypical							
L Middle Frontal G	47	0.01	15	-49	40	-6	4.08
L Middle Frontal G	10	0.01		-41	51	-7	3.53
L Inferior Parietal L	40	0.01	36	-41	-62	43	3.87
L Middle Temporal G	21	0.01	41	-62	-37	-7	4.28
L Middle Temporal G	22	0.01		-58	-39	7	3.87
L Precentral G	4	0.04	13	-31	-19	41	3.81
L Cuneus	18	0.01	574	-13	-77	26	4.46
L Precuneus	19	0.01		-6	-83	44	4.13
R Superior Frontal G	8	0.01	67	25	32	43	4.80
R Middle Frontal G	11	0.01	12	31	51	-11	3.82
R Inferior Frontal G	47	0.01	25	44	15	-5	4.14
R Inferior Parietal L	40	0.01	203	47	-46	47	4.40
R Parahippocampal G	19	0.01	73	19	-52	-2	4.83
R Posterior Cingulate	30	0.01		22	-48	12	3.69
R Postcentral G	3	0.01	16	41	-22	41	3.90
R Thalamus		0.01	13	16	-27	1	3.91
Chinese Atypical> English Atypical							
(none)							
English Typical> Chinese Typical							
L Middle Temporal G	22	0.04	36	-58	-39	7	4.06
R Inferior Frontal G	47	0.04	23	52	18	-1	3.50
R Inferior Parietal L	40	0.04	31	34	-52	42	3.98
R Middle Temporal G	22	0.04	14	52	-40	2	3.57
R Middle Temporal G	21	0.04		62	-40	-6	3.43
R Postcentral G	3	0.03	21	38	-22	41	4.74
R Precuneus	19	0.03	416	13	-81	39	4.63
L Cuneus	19	0.04		-9	-87	31	4.38
L Lingual G	19	0.04		-19	-61	-1	4.30
R Posterior Cingulate	30	0.04	45	25	-64	8	4.23
R Parahippocampal G	19	0.04		19	-49	-2	3.62
R Lingual G	19	0.04		13	-61	-1	3.35
Chinese Typical> English Typical							
(none)							

Table 3 *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)*	voxel	*x*	*y*	*z*	*Z*
English Yes > English 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 > Chinese 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

Table 4 *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 > Morphological 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> Nontransparent 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
Chinese							
Morphological Atypical> Morphological Typical							
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> Orthographic 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

Table 5 *Accuracy, reaction time and typicality rating data of Study 4.*

	Orthogrphical		**Morphological**	
	Typical	**Atypical**	**Typical**	**Atypical**
Accuracy	94.48 ± 4.68	92.30 ± 6.37	94.35 ± 6.49	94.61 ± 5.35
Reaction Time	1062.17 ± 237.12	1112.94 ± 233.92	1057.37 ± 238.16	1087.93 ± 239.14
Typicality Rating	5.05 ± 0.69	3.23 ± 0.73	5.70 ± 0.28	3.94 ± 1.07

Table 6 *Descriptive statistics on children's performance in the measures of the morphological and orthographical awareness in Study 4.*

	Mean	SD	N
Morphological compounding production (Morp_Compound)	83.69	13.6	21
Orthographical radical choice (Orth_Radical)	39.00	1.26	21
Orthographical semantic category (Orth_Semantic)	30.29	2.26	21

Table 7 *Zero-order (Pearson) correlations among all measures (N = 21).*

	1	2	3	4	5	6	7	8	9
1. **Morp_Compound**	-								
2. **Orth_Radical**	-.18	-							
3. **Orth_Semantic**	-.05	.07	-						
4. **Morp_Rating (A-T)**	.06	-.32	-.14	-					
5. **Orth_Rating(A-T)**	-.18	-.38#	.07	.70**	-				
6. **Morp_RT(A-T)**	-.33	.24	.04	.23	.26	-			
7. **Orth_RT(A-T)**	-.16	.05	-.60**	-.09	.11	.00	-		
8. **Morp_ERP_F8 (A-T)**	-.43#	.34	.09	-.04	.07	.50*	.09	-	
9. **Orth_ERP_F7 (T-A)**	.15	.15	-.34	-.12	-.02	.13	.38#	.20	-

# $p < 0.10$; * $p < 0.05$; ** $p < 0.01$

Appendices

Appendix 1 *Category-level typicality ratings (1-6) in Pilot Study 2 show similarities across languages for Typical vs. Atypical items for the ten categories used in Study 1.* Typical items (Chinese, $M = 5.13$, $SD = 0.39$; English, $M = 5.74$, $SD = 0.19$), Atypical items (Chinese, $M = 3.70$, $SD = 1.13$; English, $M = 3.90$, $SD = 0.63$). A Typicality (Typical vs. Atypical) by Language (English vs. Chinese) ANOVA revealed only a main effect of Typicality, $F(1, 18) = 53.68$, $P < 0.001$. No significant effects of Language or Typicality by Language interactions were found.

	Chinese			
	Typical		**Atypical**	
SHOES	loafers	5.33	slippers	4.30
PANTS	trousers	5.38	overalls	4.83
VEHICLE	car	5.52	train	4.21
WRITING INSTRUMENT	pencil	5.00	chalk	3.17
STATION	train station	4.74	airport	1.96
BUG	fly	4.74	butterfly	4.08
BUILDING	office building	5.75	garage	1.50
OIL	engine oil	4.39	gasoline	3.41
VEGETABLE	celery	5.30	eggplant	4.74
PAPER	writing paper	5.13	toilet paper	4.87
	English			
	Typical		**Atypical**	
SHOES	loafers	5.30	slippers	3.15
PANTS	trousers	5.85	overalls	4.19
VEHICLE	car	5.96	train	3.59
WRITING INSTRUMENT	pencil	5.81	chalk	4.33
STATION	train station	5.77	airport	3.59
BUG	fly	5.85	butterfly	4.74
BUILDING	office building	5.85	garage	2.84
OIL	engine oil	5.54	gasoline	3.74
VEGETABLE	celery	5.59	eggplant	4.96
PAPER	writing paper	5.89	toilet paper	3.85

Appendix 2 *Category-level typicality ratings (1-6) from Pilot Study 3 show similarities across Label types for Typical vs. Atypical exemplar pictures for the ten categories used in study 2.* Typical items (Morphological, *M* = 5.34, *SD* = 0.14; Orthographic, *M* = 5.09, *SD* = 0.32), Atypical items (Morphological, *M* = 4.28, *SD* = 0.69; Orthographic, *M* = 3.88, *SD* = 0.76). A Typicality (Typical vs. Atypical) by Label type (Morphological vs. Orthographic) ANOVA revealed only a main effect of Typicality, *F* (1, 4) = 22.97, *P* = 0.009. No main effect of Label type or interactions between Typicality and Label type were found.

	Morphological			
	Typical		**Atypical**	
SHOES *xie2zi*	loafers *pi2xie2*	5.33	slippers *tuo1xie2*	4.30
PANTS *ku4zi*	trousers *xi1ku4*	5.38	overalls *bei1dai4ku4*	4.83
VEHICLE *che1*	car *jiao4che1*	5.52	train *huo3che1*	4.21
WRITING INSTRUMENT *bi3*	pencil *qian1bi3*	5.38	chalk *fen3bi3*	3.17
PAPER *zhi3*	writing paper *xin4zhi3*	5.13	toilet paper *shou3zhi3*	4.87
	Orthographic			
VEGETABLE *cai4*	celery *xi1qin2*	5.30	eggplant *qie2zi3*	4.74
BUG *chong2zi*	fly *cang1ying*	4.74	butterfly *hu2die2*	4.08
BIRD *niao3*	pigeon *ge1zi3*	5.43	penguin *qi3e2*	3.00
SHIP *chuan2*	warship *jun1jian4*	5.22	yacht *you2ting3*	4.39
STONE *shi2tou2*	rock *yan2shi*	4.74	brick *zhuan1tou2*	3.17

Appendix 3 *Category-level typicality ratings (1-6) from Pilot Study 5 show similarities across Label types for Typical vs. Atypical exemplar pictures for the ten categories used in study 3.* Typical items (Morphological, *M* = 5.45, *SD* = 0.07; Nontransparent, *M* = 5.67, SD = 0.09), Atypical items (Morphological, *M* = 3.67, *SD* = 0.15; Nontransparent, *M* = 4.03, *SD* = 0.18). A Typicality (Typical vs. Atypical) by Label type (Morphological vs. Nontransparent) ANOVA revealed only a main effect of Typicality, *F* (1, 4) = 35.60, *P* = 0.004. No significant main effect of Label type or interactions between Typicality and Label type were found.

	Morphological			
	Typical		**Atypical**	
PAPER	writing paper	5.13	toilet paper	4.87
PHONE	cell phone	5.76	rotary phone	3.68
BALL	basketball	5.55	football	4.48
BOOK	textbook	5.13	notebook	3.48
CHAIR	folding chair	5.48	rocking chair	3.24
	Nontransparent			
	Typical		**Atypical**	
SHOES	loafers	5.30	slippers	3.15
PANTS	trousers	5.85	overalls	4.19
VEHICLE	car	5.96	train	3.59
WRITING INSTRUMENT	pencil	5.81	chalk	4.33
VEGETABLE	celery	5.59	eggplant	4.96

Appendix 4 *Category-level object picture typicality rating (1-6) results for the six categories used in cross-linguistic comparison.*

Language	Category	Item	Mean	Std. E
English	Vegetable	Eggplant	3.33	.29
		Celery	4.33	.24
	Vehicle	Train	3.06	.29
		Car	5.72	.12
	Writing Instrument	Chalk	2.94	.28
		Pencil	5.61	.12
	Paper	Toilet paper	2.72	.31
		Writing paper	5.06	.20
	Pants	Overall	2.44	.23
		Trousers	5.56	.09
	Shoes	Slipper	2.28	.24
		Loafer	5.00	.16
Chinese	Vegetable*cai4*	Eggplant *qie2zi*	3.67	.29
		Celery *xi1qin2*	5.50	.24
	Vehicle *che1*	Train *huo3che1*	3.78	.29
		Car *jiao4che1*	5.78	.12
	Writing Instrument*bi3*	Chalk *fen3bi3*	2.78	.28
		Pencil *qian1bi3*	5.67	.12
	Paper *zhi3*	Toilet paper *shou3zhi3*	3.61	.31
		Writing paper *xin4zhi3*	5.56	.20
	Pants *ku4zi*	Overall *bei1dai4ku4*	4.06	.23
		Trousers *xi1ku4*	5.94	.09
	Shoes *xie2zi*	Slippers *tuo1xie2*	3.89	.24
		Loafers *pi2xie2*	5.83	.16

Appendix 5 *Category-level object picture typicality rating (1-6) results for the ten categories used in English participants.*

Label	Category	Item	Mean	Std. E
Nontransparent	Vegetable	Eggplant	3.33	.32
		Celery	4.33	.30
	Vehicle	Train	3.06	.26
		Car	5.72	.14
	Writing Instrument	Chalk	2.94	.25
		Pencil	5.61	.12
	Pants	Overall	2.44	.23
		Trousers	5.56	.12
	Shoes	Slippers	2.28	.24
		Loafers	5.00	.21
Morphological	Chair	Rocking Chair	2.89	.25
		Folding Chair	4.83	.23
	Phone	Rotary Phone	3.44	.42
	_	Cell Phone	5.06	.29
	Ball	Football	3.33	.31
	_	Basketball	5.72	.11
	Paper	Toilet paper	2.72	.30
	_	Writing paper	5.06	.22
	Book	Notebook	2.44	.23
		Textbook	5.28	.16

Appendix 6 *Category-level object picture typicality rating (1-6) results for the ten categories used in Chinese participants.*

Label	Category	Item	Mean	Std. E
Orthographic	**Bird***niao3*	**Penguin***qi3e2*	2.83	.29
		Pigeon*ge1zi3*	5.56	.19
	Vegetable*cai4*	**Eggplant** *qie2zi3*	3.67	.26
		Celery *xi1qin2*	5.50	.17
	Stone*shi2tou2*	**Brick***zhuan1tou2*	2.33	.28
		Rock*yan2shi*	5.72	.14
	Ship*chuan2*	**Yacht***you2ting3*	4.17	.27
		Warship*jun1jian4*	5.39	.24
	Bug *chong2zi*	**Butterfly** *hu2die2*	2.94	.26
		Fly *cang1ying*	4.94	.26
Morphological	**Vehicle** *che1*	**Train** *huo3che1*	3.78	.32
		Car *jiao4che1*	5.78	.10
	Writing Instrument*bi3*	**Chalk** *fen3bi3*	2.78	.31
		Pencil *qian1bi3*	5.67	.11
	Paper *zhi3*	**Toilet paper** *shou3zhi3*	3.61	.32
		Writing paper *xin4zhi3*	5.56	.17
	Pants *ku4zi*	**Overall** *bei1dai4ku4*	4.06	.24
		Trousers *xi1ku4*	5.94	.06
	Shoes *xie2zi*	**Slippers** *tuo1xie2*	3.89	.24
		Loafers *pi2xie2*	5.83	.09

Appendix 7. *Chinese morphological compounding production task materials* (P. D. Liu & McBride-Chang, 2010).

词素意识测试 - II

指导语:上面那些题目都给你提供了一个词语作为提示,下面还有一些题目我将不再提供提示，请你以此类推，根据所提的问题，发挥你的想象力，给出一个你认为最合适、最能表达问题中意思的词语。这些词语可以是现实生活中不会用到的词。下面我还是给你几个例子来学习一下：

我们把红色的星星叫做什么呢？ ------------------------------------ 红星
我们把能喷出烟的洞穴叫什么？ ------------------------------------ 喷烟洞
我们把用草制成的伞叫做什么呢? ------------------------------------ 草伞
我们把大象的吼叫称为什么? ------------------------------------ 象吼
我们把收集气体叫做什么? ------------------------------------ 集气
车子和桌子放在一起我们可以叫做什么? ------------------------------------ 车桌

	题目	儿童答案	标准答案	评分
26	能够把人缩小但不能缩小别的东西的枪, 我们叫它什么呢？	________	缩人枪	________
24	我们把专门吃铁块的怪兽叫做什么?	________	吃铁怪	________
63	云彩和鲜花放在一起可以叫做什么？	________	云花	________
42	我们把星星的闪烁叫做什么?	________	星闪	________
32	我们把像海绵一样的橡皮叫做什么？	________	海绵橡皮	________
35	味道像菠萝的桃子, 我们把它叫做什么呢?	________	菠萝桃	________
14	我们把味道酸酸的雾气叫做什么？	________	酸雾	________
44	我们把叶子的生长叫做什么?	________	叶长	________
45	我们把车辆的摆动叫做什么？	________	车摆	________
33	外星上有种树能够长出鸡蛋，我们应该把这种树叫做什么呢？	________	鸡蛋树	________
43	我们把泥土的流动叫什么呢？	________	泥流	________
55	把金子冷冻起来叫做什么?	________	冻金	________
62	木头和铁片放在一起可以叫做什么？	________	木铁	________
51	把水晶从地下挖掘出来叫做什么	________	挖水晶	________
21	我们把专门用来切石头的刀叫什么？	________	切石刀	________
11	我们把稀疏的森林叫做什么?	________	稀 林（ 疏林）	________

IRB: «IRB» IRB Number: «ID» **Document Approved On: «ApprovalDate»**

Page 1 of 2

34	有一种样子长的像肉一样的石头，我们把这种石头叫做什么呢？	________	肉石	________
65	水稻和小麦放在一起可以叫做什么？	________	稻麦	________
25	我们应该如何称呼专门从事修桥工作的人？	________	修桥工	________
54	把字从书本上扫掉叫做什么？	________	扫字	________
13	我们把温暖的夜色叫做什么？	________	暖夜	________
41	我们把小马在跳舞叫做什么？	________	马舞	________
64	小鸟、螃蟹、蜘蛛放在一起可以叫做什么？	________	鸟蟹蛛	________
53	我们把对轮船进行清洗叫做什么？	________	洗船	________
22	我们把专门用来碾碎玉米的机器叫做什么？	________	碾玉米机	________
31	我们把用雪建成的房子叫做什么？	________	雪房	________
23	我们把只能显示数字的屏幕叫什么？	________	数屏	________
15	我们把非常柔软的铁叫做什么？	________	软铁	________
52	我们把对着眼睛吹气叫做什么？	________	吹眼	________
61	房子、树木和小猪放在一起可以叫做什么？	________	房树猪	________
12	我们把喜庆的音乐叫做什么？	________	喜乐	________

评分标准：

结构正确、简洁，使用了所有的关键语素，正确且完整地表达了句子中的意思 -- 4 分

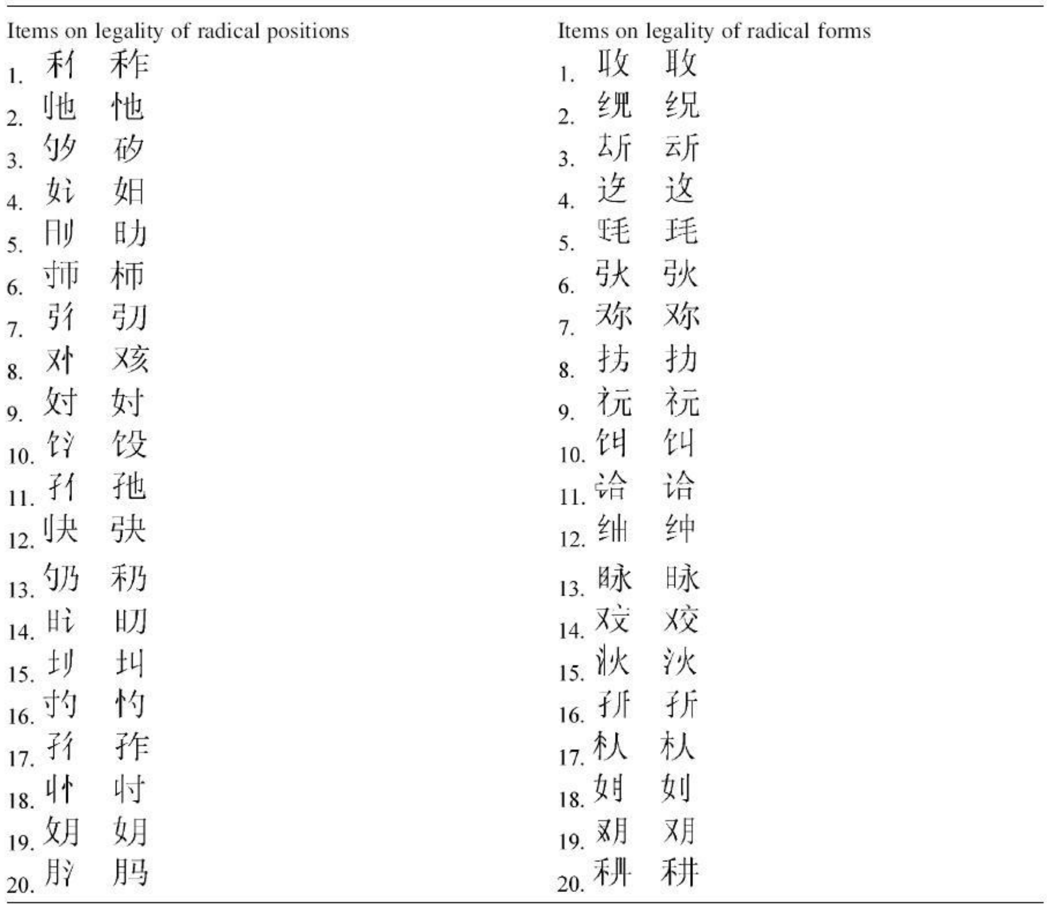
结构正确，但冗长，使用了句子中的非关键语素（但没有简单重复）或使用了语义上与关键语素相近的语素，能够正确完整地表达句子中的意思 -- 3 分

结构正确但冗长或者不完整（即缺少关键语素，但该语素对词语整体结构未造成影响），或出现无关语素，或简单重复句子中的内容，不能完整表达句子中的意思 -- 2 分

结构不正确（如结构颠倒，或缺少决定词语结构的关键语素），但使用了其他关键语素或相关语素，与整句意思有一定的关系 -- 1 分

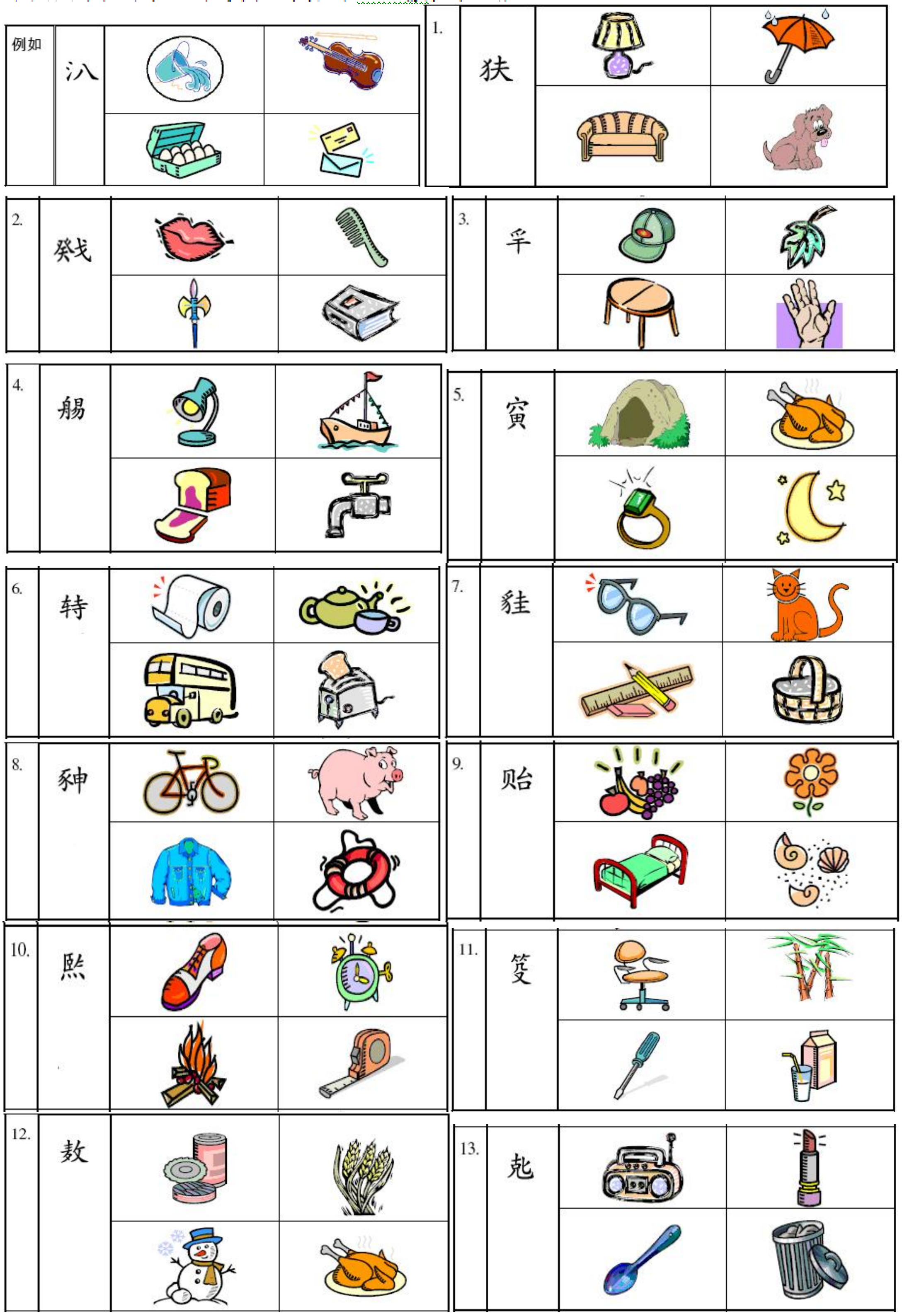
结构不正确，使用了无关语素，与整句意思的相关性不高 -- 0 分

Appendix 8. *Chinese Orthographic radical choice task materials (Wang et al., 2005).*

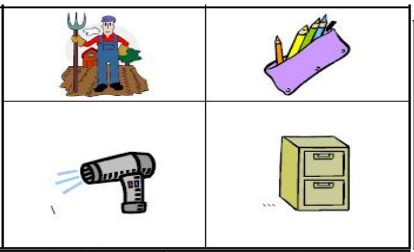


Appendix 9. *Chinese orthographical semantic category task materials* (Tong, 2008; Tong & McBride-Chang, 2009).

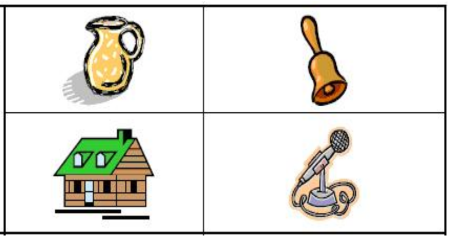
下面四个图中哪一个最好的代表了左边假汉字可能的意思?



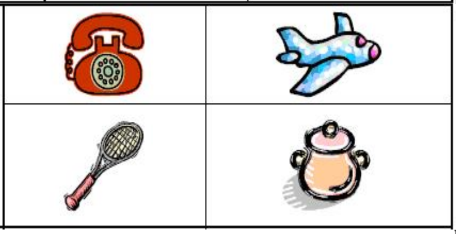
14. 㨃



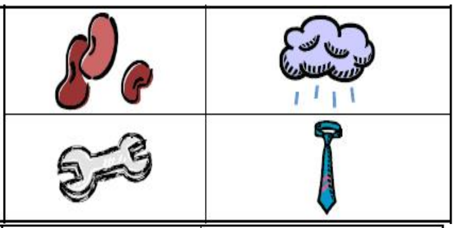
15. ⿸广者



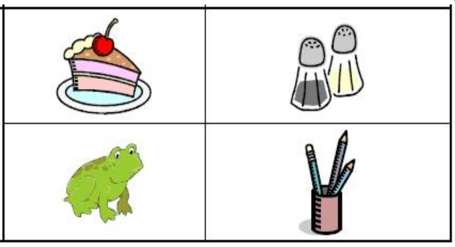
16. ⿱臣皿



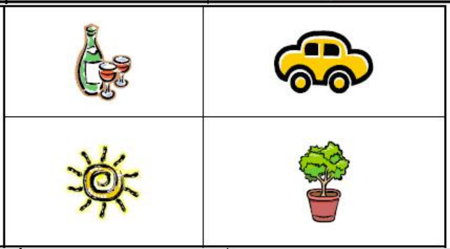
17. ⿱次豆



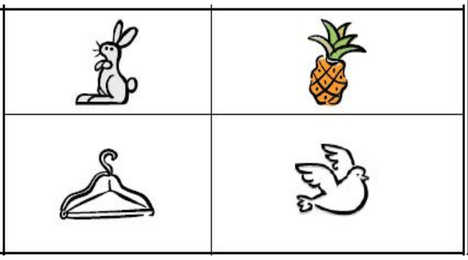
18. ⿱般龟



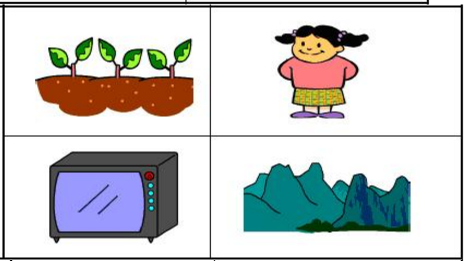
19. ⿰酉主



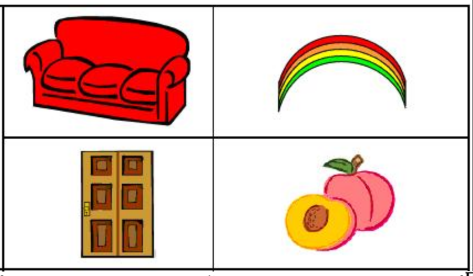
20. ⿰米鸟



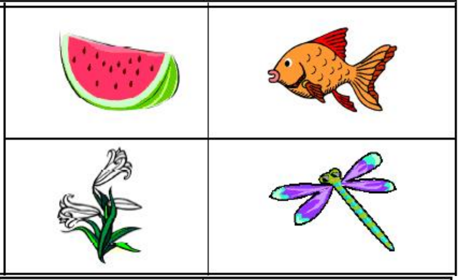
21. ⿰禾㐬



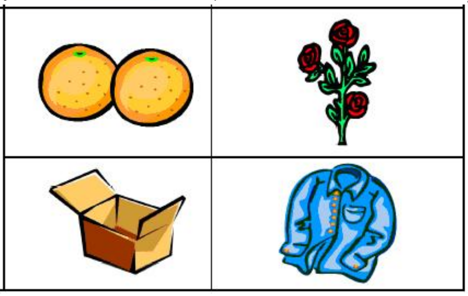
22. ⿵门国



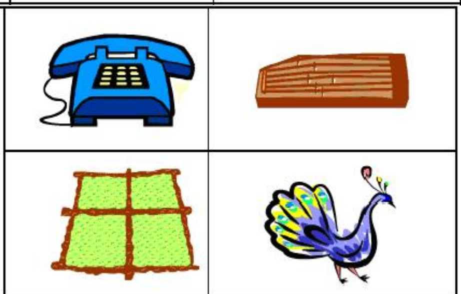
23. ⿰鱼央



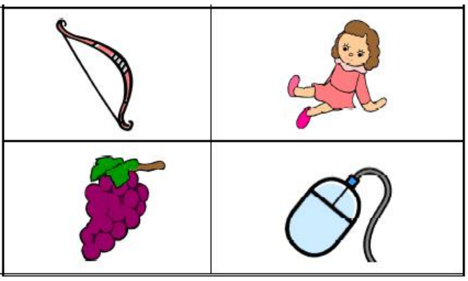
24. ⿱次衣



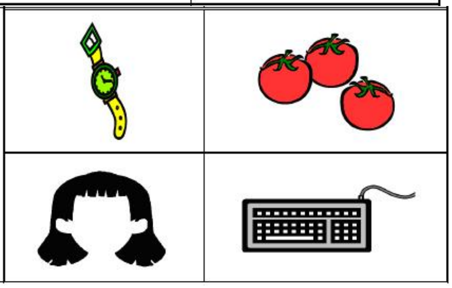
25. ⿰目豕



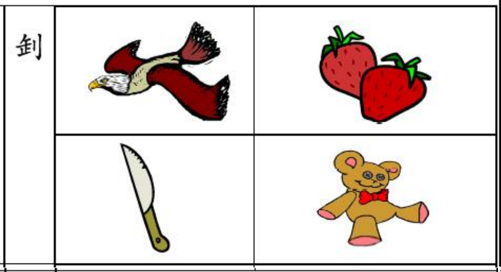
26. ⿰弓勺



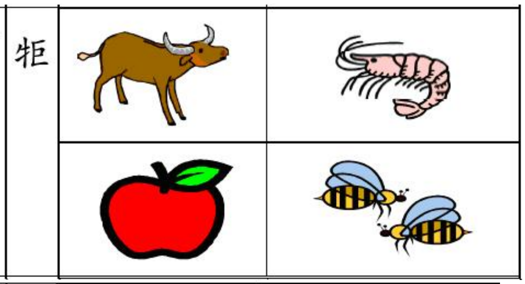
27. ⿱夸廾



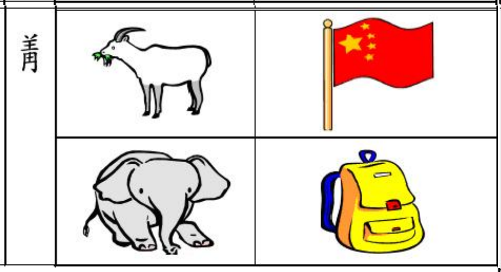
28.



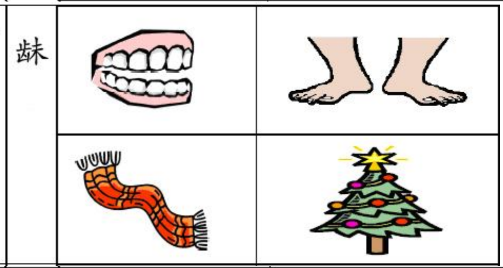
29.



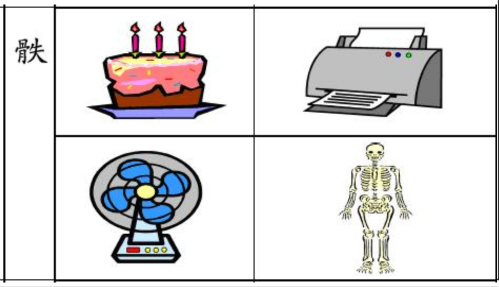
30.



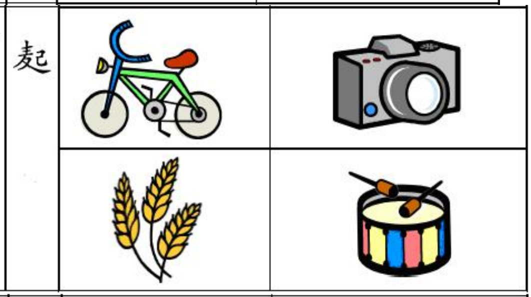
31.



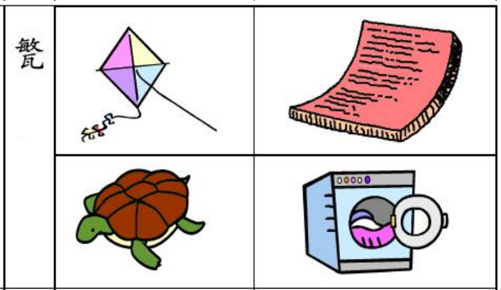
32.



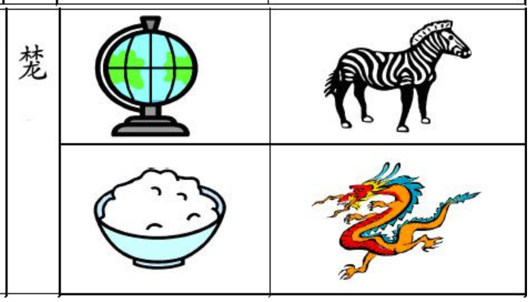
33.



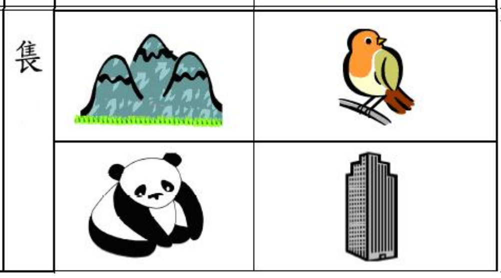
4.



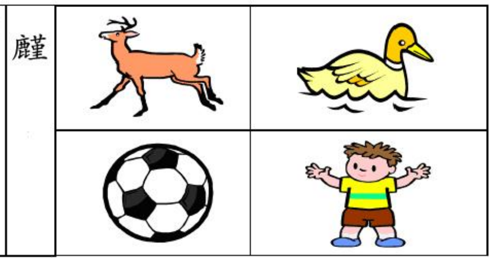
35.



36.



37.



1. Tong, X., & McBride-Chang, C. (2009). *Development of Chinese word reading: Changing structural patterns of metalinguistic awareness and other cognitive skills.* Paper posted at the 2009 SRCD Biennial Meeting Biennial Meeting of Society for Research in Child Development. Denver, Colorado. (April 2).

2. Tong, X. (2008).The Development of Chinese Word Reading: Relations of Sub-Character Processing, Phonological Awareness, Morphological Awareness, and Orthographic Knowledge to Chinese-English Biscriptal Reading. CUHK dissertation.

Trial		A		10		C		20		D		30		A	
1		D		11		B		21		A		31		A	
2		C		12		B		22		C		32		D	
3		D		13		C		23		B		33		C	
4		B		14		A		24		D		34		B	
5		A		15		C		25		C		35		D	
6		C		16		D		26		A		36		B	
7		B		17		A		27		C		37		A	
8		B		18		C		28		C					
9		D		19		A		29		A					

References

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

Anglin, J. (1977). *Word, object, and conceptual development*. New York: Norton.

Ashby, F. G., & Casale, M. B. (2005). Empirical dissociations between rule-based and similarity-based categorization. *Behavioral and Brain Sciences, 28*(1), 15-+.

Bailenson, J. N., Shum, M. S., Atran, S., Medin, D. L., & Coley, J. D. (2002). A bird's eye view: biological categorization and reasoning within and across cultures. *Cognition, 84*(1), 1-53.

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

Basso, A., Capitani, E., & Laiacona, M. (1988). Progressive language impairment without dementia - a case with isolated category specific semantic defect. *Journal of Neurology Neurosurgery and Psychiatry, 51*(9), 1201-1207.

Batty, M., & Taylor, M. J. (2002). Visual categorization during childhood: An ERP study. *Psychophysiology, 39*(4), 482-490.

Bjorklund, D. F., Thompson, B. E., & Ornstein, P. A. (1983). Developmental-Trends in Childrens Typicality Judgments. *Behavior Research Methods & Instrumentation, 15*(3), 350-356.

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

Boster, J. S. (1988). Natural Sources of Internal Category Structure - Typicality, Familiarity, and Similarity of Birds. *Memory & Cognition, 16*(3), 258-270.

Brett, M., Anton, J.-L., Valabregue, R., & Poline, J.-B. (2002). *Region of interest analysis using an SPM toolbox.* 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.

Chen, X., Hao, M. L., Geva, E., Zhu, J., & Shu, H. (2009). The role of compound awareness in Chinese children's vocabulary acquisition and character reading. *Reading and Writing, 22*(5), 615-631.

Chiu, L. H. (1972). Cross-Cultural Comparison of Cognitive Styles in Chinese and American Children. *International Journal of Psychology, 7*(4), 235-242.

Chumbley, J. I. (1986). The Roles of Typicality, Instance Dominance, and Category Dominance in Verifying Category Membership. *Journal of Experimental Psychology-Learning Memory and Cognition, 12*(2), 257-267.

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.

Diesendruck, G., & haLevi, H. (2006). The role of language, appearance, and culture in children's social category-based induction. *Child Development, 77*(3), 539-553.

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.

Gelman, S. A., & Markman, E. M. (1987). Young Childrens Inductions from Natural Kinds - the Role of Categories and Appearances. *Child Development, 58*(6), 1532-1541.

Gelman, S. A., Wilcox, S. A., & Clark, E. V. (1989). Conceptual and Lexical Hierarchies in Young-Children. *Cognitive Development, 4*(4), 309-326.

Gratton, G., Coles, M. G. H., & Donchin, E. (1983). A New Method for Off-Line Removal of Ocular Artifact. *Electroencephalography and Clinical Neurophysiology, 55*(4), 468-484.

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., Koenig, P., Glosser, G., DeVita, C., Moore, P., Rhee, J., et al. (2003). Neural basis for semantic memory difficulty in Alzheimer's disease: an fMRI study. *Brain, 126*, 292-311.

Grossman, M., Koenig, P., Kounios, J., McMillan, C., Work, M., & Moore, P. (2006). Category-speciric effects in semantic memory: Category-task interactions suggested by fMRI. *Neuroimage, 30*(3), 1003-1009.

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

Han, S. H., Jiang, Y., Gu, H., Rao, H. Y., Mao, L. H., Cui, Y., et al. (2004). The role of human parietal cortex in attention networks. *Brain, 127*, 650-659.

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.

Hauk, O., Patterson, K., Woollams, A., Watling, L., Pulvermuller, F., & Rogers, T. T. (2006). [Q:] When would you prefer a SOSSAGE to a SAUSAGE? [A:] At about 100 msec. ERP correlates of orthographic typicality and lexicality in written word recognition. *Journal of Cognitive Neuroscience, 18*(5), 818-832.

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.

Imai, M., & Haryu, E. (2001). Learning proper nouns and common nouns without clues from syntax. *Child Development, 72*(3), 787-802.

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.

Junghofer, M., Elbert, T., Tucker, D. M., & Braun, C. (1999). The polar average reference effect: a bias in estimating the head surface integral in EEG recording. *Clinical Neurophysiology, 110*(6), 1149-1155.

Kiefer, M. (2001). Perceptual and semantic sources of category-specific effects: Event-related potentials during picture and word categorization. *Memory & Cognition, 29*(1), 100-116.

Knibb, J. A., & Hodges, J. R. (2005). Semantic dementia and primary progressive aphasia: A problem of categorization? *Alzheimer Disease & Associated Disorders, 19*, S7-S14.

Koenig, P., Smith, E. E., DeVita, C., Moore, P., McMillan, C., & Grossman, M. (2002). Rule- and similarity-based acquisition of a novel category: An fMRI study. *Journal of Cognitive Neuroscience*, B41.

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.

Koenig, P., Smith, E. E., Moore, P., Glosser, G., & Grossman, M. (2007). Categorization of novel animals by patients with Alzheimer's disease and corticobasal degeneration. *Neuropsychology, 21*(2), 193-206.

Koenig, P., Smith, E. E., Troiani, V., Anderson, C., Moore, P., & Grossman, M. (2008). Medial Temporal Lobe Involvement in an Implicit Memory Task: Evidence of Collaborating Implicit and Explicit Memory Systems from fMRI and Alzheimer's Disease. *Cerebral Cortex, 18*(12), 2831-2843.

Komatsu, L. K. (1992). Recent Views of Conceptual Structure. *Psychological Bulletin, 112*(3), 500-526.

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.

Lang, M., Lang, W., Uhl, F., Kornhuber, A., Deecke, L., & Kornhuber, H. H. (1987). SLOW NEGATIVE POTENTIAL SHIFTS INDICATING VERBAL COGNITIVE-LEARNING IN A CONCEPT-FORMATION TASK. *Human Neurobiology, 6*(3), 183-190.

Lin, P. J., & Schwanenflugel, P. J. (1995). Cultural Familiarity and Language Factors in the Structure of Category Knowledge. *Journal of Cross-Cultural Psychology, 26*(2), 153-168.

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

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.

Liu, D., Meltzoff, A. N., & Wellman, H. M. (2009). Neural Correlates of Belief- and Desire-Reasoning. *Child Development, 80*(4), 1163-1171.

Liu, D., Sabbagh, M. A., Gehring, W. J., & Wellman, H. M. (2009). Neural Correlates of Children's Theory of Mind Development. *Child Development, 80*(2), 318-326.

Liu, P. D., & McBride-Chang, C. (2010). What Is Morphological Awareness? Tapping Lexical Compounding Awareness in Chinese Third Graders. *Journal of Educational Psychology, 102*(1), 62-73.

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. M., & Shatz, M. (1996). Linguistic influences on categorization in preschool children: A crosslinguistic study. *Journal of Child Language, 23*(3), 529-545.

McBride-Chang, C., Cho, J. R., Liu, H. Y., Wagner, R. K., Shu, H., Zhou, A. B., et al. (2005). Changing models across cultures: Associations of phonological awareness and morphological structure awareness with vocabulary and word recognition in second graders from Beijing, Hong Kong, Korea, and the United States. *Journal of Experimental Child Psychology, 92*(2), 140-160.

McClelland, J. L., Rogers, T. T., Patterson, K., Dilkina, K. N., & Lambon Ralph, M. R. (2009). Semantic Cognition: Its Nature, Its Development, and its Neural Basis. In M. Gazzaniga (Ed.), *The Cognitive Neurosciences IV*. Boston, MA: MIT Press.

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., & Pani, J. R. (1980). Acquisition of Basic Object Categories. *Cognitive Psychology, 12*(4), 496-522.

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.

Nisbett, R. E., & Miyamoto, Y. (2005). The influence of culture: holistic versus analytic perception. *Trends in Cognitive Sciences, 9*(10), 467-473.

Patalano, A. L., Smith, E. E., Jonides, J., & Koeppe, R. A. (2001). PET evidence for multiple strategies of categorization. *Cognitive, Affective & Behavioral Neuroscience, 1*(4), 360-370.

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.
Poldrack, R. A., Wagner, A. D., Prull, M. W., Desmond, J. E., Glover, G. H., & Gabrieli, J. D. E. (1999). Functional specialization for semantic and phonological processing in the left inferior prefrontal cortex. *Neuroimage, 10*(1), 15-35.
Quinn, P. C., Westerlund, A., & Nelson, C. A. (2006). Neural markers of categorization in 6-month-old infants. *Psychological Science, 17*(1), 59-66.
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. (2004). *Semantic Cognition: A Parallel Distributed Processing Approach.* Cambridge, MA: MIT Press.
Rogers, T. T., & McClelland, J. L. (2008). Precis of Semantic Cognition: A Parallel Distributed Processing Approach. *Behavioral and Brain Sciences, 31*(6), 689-+.
Rogers, T. T., & McClelland, J. L. (2010). Semantics without categorization. In E. M. Pothos & A. J. Wills (Eds.), *Formal approaches to categorization*. Cambridge, UK: Cambridge University Press.
Rogers, T. T., Ralph, M. A. L., Garrard, P., Bozeat, S., McClelland, J. L., Hodges, J. R., et al. (2004). Structure and deterioration of semantic memory: A neuropsychological and computational investigation. *Psychological Review, 111*(1), 205-235.
Rosch, E. H. (1973). Natural Categories. *Cognitive Psychology, 4*(3), 328-350.
Russell, J. A. (1991). Culture and the Categorization of Emotions. *Psychological Bulletin, 110*(3), 426-450.
Schunn, C. D., & Vera, A. H. (2004). Cross-cultural similarities in category structure. *Thinking & Reasoning, 10*(3), 273-287.
Schwanenflugel, P. J., & Rey, M. (1986). The Relationship between Category Typicality and Concept Familiarity - Evidence from Spanish-Speaking and English-Speaking Monolinguals. *Memory & Cognition, 14*(2), 150-163.
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., Jin, Z., Fletcher, P., & Tan, L. H. (2003). Distinct brain regions associated with syllable and phoneme. *Human Brain Mapping, 18*(3), 201-207.
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.
Smith, E. E., & Medin, D. L. (1981). *Categories and Concepts*. Cambridge: Harvard University Press
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.
Tan, L. H., Laird, A. R., Li, K., & Fox, P. T. (2005). Neuroanatomical correlates of phonological processing of Chinese characters and alphabetic words: A meta-analysis. *Human Brain Mapping, 25*(1), 83-91.
Tardif, T. (2006). *But are They Really Verbs?: Chinese Words for Action*: Oxford.
Tong, X. (2008). *The Development of Chinese Word Reading: Relations of Sub-Character Processing, Phonological Awareness, Morphological Awareness, and Orthographic Knowledge to Chinese-English Biscriptal Reading.* . Unpublished Dissertation, Chinese University of Hong Kong, Disstertation, Hong Kong.
Tong, X., & McBride-Chang, C. (2009). *Development of Chinese word reading: Changing structural patterns of metalinguistic awareness and other cognitive skills.* Paper presented at the 2009 SRCD Biennial Meeting Biennial Meeting of Society for Research in Child Development.,
Treiman, R., & Cassar, M. (1997). Can children and adults focus on sound as opposed to spelling in a phoneme counting task? *Developmental Psychology, 33*(5), 771-780.
Unsworth, S. J., Sears, C. R., & Pexman, P. M. (2005). Cultural influences on categorization processes. *Journal of Cross-Cultural Psychology, 36*(6), 662-688.
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.
Wang, M., Cheng, C. X., & Chen, S. W. (2006). Contribution of morphological awareness to Chinese-English biliteracy acquisition. *Journal of Educational Psychology, 98*(3), 542-553.
Wang, M., Perfetti, C. A., & Liu, Y. (2005). Chinese-English biliteracy acquisition: cross-language and writing system transfer. *Cognition, 97*(1), 67-88.
Waxman, S. R., & Markow, D. B. (1995). Words as invitations to form categories: Evidence from 12- to 13-month-old infants. *Cognitive Psychology, 29*(3), 257-302.
Whorf, B. L. (1956). *Language, Thought and Reality: Selected Writings of Benjamin Lee Whorf*. Cambridge,MA: MIT Press.
Woollams, A. M., Cooper-Pye, E., Hodges, J. R., & Patterson, K. (2008). Anomia: A doubly typical signature of semantic dementia. *Neuropsychologia, 46*(10), 2503-2514.
Yoshida, H., & Smith, L. B. (2003). Shifting ontological boundaries: how Japanese- and English-speaking children generalize names for animals and artifacts. *Developmental Science, 6*(1), 1-17.
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.
Zhou, Y. G. (1978). To what degree are the "phonetics" of present-day Chinese characters still phonetic? *Zhongguo Yuwen, 146*, 172-177.