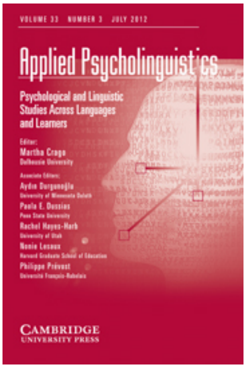


# Cognitive mechanism of writing to dictation of logographic characters

ZAIZHU HAN, LUPING SONG and YANCHAO BI

Applied Psycholinguistics / Volume 33 / Issue 03 / July 2012, pp 517 - 537
DOI: 10.1017/S0142716411000464, Published online: 08 August 2011

Applied Psycholinguistics **33** (2012), 517–537
doi:10.1017/S0142716411000464

# Cognitive mechanism of writing to dictation of logographic characters

ZAIZHU HAN
*Beijing Normal University*

LUPING SONG
*Capital Medical University Rehabilitation College, China*

YANCHAO BI
*Beijing Normal University*

Received: September 10, 2009 Accepted for publication: October 24, 2010

ADDRESS FOR CORRESPONDENCE
Yanchao Bi, National Key Laboratory of Cognitive Neuroscience and Learning, Beijing Normal University, Beijing 100875, People's Republic of China. E-mail: ybi@bnu.edu.cn

ABSTRACT
The cognitive mechanisms for writing to dictation of Chinese syllables by healthy adults were investigated using large-sample multiple regression analyses. In the experiment, subjects wrote down a corresponding character upon hearing a syllable. We mainly examined the effects of three types of attributes (i.e., lexical, semantic, and phonology to orthography conversion [POC] ones) in predicting the production probability of specific characters out of the homophone families for target syllables. We observed significant effects for all three types of attributes, as well as interactions between POC and the lexical attributes, and between POC and the semantic attributes. We further found that the semantic effects vanished for the writing stimuli without homophones. A feedback procedure (i.e., phonetic radical transparency) was also observed to influence Chinese writing performances. Our results support the hypothesis that the extent of semantic involvement in writing (spelling) to dictation is influenced by the effectiveness of POC procedure in a certain language and/or word set. The existence of an interaction between the lexical semantic route and the POC route in writing is further consolidated.

Writing to dictation, the process of producing orthographic output given an auditory input, is often viewed as the counterpart of reading. It is generally accepted that writing to dictation involves three potential anatomically and functionally distinctive processing routes (e.g., Norton, Kovelman, & Petitto, 2007). One is by converting the phonological input directly into the orthographic output according to nonlexical phonology to orthography conversion (POC) rules (e.g., Luzzi et al., 2003). In alphabetic scripts such as English, the POC rules are realized by phoneme–grapheme mapping (e.g., /f/ might map onto "f," "ph," or "gh"). The second route is mediated by the lexical semantic system, going through the phonological (input) lexicon, the semantic system, and the orthographic (output) lexicon (e.g., Bub & Kertesz, 1982; Hillis & Caramazza, 1991, 1995; Miceli, Benvegnu, Capasso, & Caramazza, 1997; Nolan & Caramazza, 1983). Finally, there is potentially a third route that matches the phonological (input) lexical representation to the orthographic (output) lexical representation, bypassing the semantic system (see Houghton & Zorzi, 2003). The POC route is competent in spelling to dictation of nonwords and words following regular POC, and is influenced by the degree of POC transparency (e.g., Delattre, Bonin, & Barry, 2006; Rapp, Epstein, & Tainturier, 2002). The two lexical routes are efficient in handling familiar words, and are sensitive to lexical factors (e.g., word frequency [WF], age of acquisition [AoA]) and semantic factors (e.g., conceptual imageability [IMG], concreteness [CON], familiarity [FAM]; e.g., Bonin, Barry, Méot, & Chalard, 2004; Bonin & Méot, 2002). These routes might interact with each other (e.g., Bosse, Valdois, & Tainturier, 2003; Delattre et al., 2006; Folk & Jones, 2004; Hillis & Caramazza, 1991, 1995; Hillis, Rapp, & Caramazza, 1999; Laiacona et al., 2009; Rapp et al., 2002; Ward, Stott, & Parkin, 2000).

This generic theory of writing to dictation has been motivated primarily by various types of impaired writing to dictation performances of brain-damaged individuals. Patients in alphabetic scripts have been reported to show advantages of spelling to dictation regular words and nonwords over irregular words, and made regularization errors (e.g., "yacht" ➔ yot), indicating heavy reliance on the POC route (*surface dysgraphia*; e.g., Baxter & Warrington, 1987; Beauvois & Dérouesné, 1981). Some other patients made semantic errors in spelling to dictation (e.g., "time" ➔ clock), indicating the involvement of the lexical semantic route (*deep dysgraphia*; e.g., Bub & Kertesz, 1982; Duhamel & Poncet, 1986; Majerus, Lekeu, Van der Linden, & Salmon, 2001; Miceli, Capasso, Ivella, & Caramazza, 1997; Tainturier & Caramazza, 1996; Valdois, Carbonnel, David, Rousset, & Pellat, 1995; Weekes & Raman, 2008). These patients are often poorer with words of low IMG/CON (i.e., less rich semantic properties). A third pattern is associated with patients who had impaired nonword spelling to dictation, relative to preserved ability to spell to dictation of real words with irregular POC, and did not make semantic errors in writing (*phonological dysgraphia;* e.g., Shallice, 1981). Similar dysgraphic symptoms have also been found in patient groups of logographic writing systems such as Chinese (*surface dysgraphia*; e.g., Law & Or, 2001; Reich, Chou, & Patterson, 2003; *deep dysgraphia*; e.g., Gao et al., 1993; *phonological dysgraphia*; e.g., Kim & Na, 2000). Evidence of interaction among these routes has also been observed in some dysgraphic individuals (e.g., Folk & Jones, 2004; Hillis & Caramazza, 1991, 1995; Hillis et al., 1999; Laiacona et al., 2009; Miceli & Capasso, 2006; Rapp et al., 2002). For example, Hillis et al. (1999) found that their patient (RCM) had deficits in the writing routes. Of interest, along with the recovery of the POC route (accuracy of pseudoword writing to dictation = 42%–67%), the semantic errors in writing to dictation dropped accordingly (56% to 11%). This pattern indicates the dynamic interaction between the POC and lexical semantic routes in constraining responses of writing to dictation. When the POC route was less impaired, it could more effectively block potential responses being generated through the lexical semantic route that did not fit the POC rules (semantically related candidates).

It is theoretically unclear, however, whether the models derived from the rich acquired dysgraphic patterns can be generalized directly to normal writing to dictation systems. It is possible that some mechanisms are only applied in pathological systems to make up for the writing deficits and play a lesser role in normal circumstances. Studies on the normal writing to dictation processes of alphabetic scripts have shown the effects of the POC component, lexical factors, and their interactions (e.g., Barry & Seymour, 1988; Bosse et al., 2003; Campbell, 1983; Delattre et al., 2006). The involvement of a lexical route has been repeatedly demonstrated in priming studies with English speakers (e.g., Barry & Seymour, 1988; Campbell, 1983). For example, Cambell (1983) asked subjects to hear a pseudoword sound (e.g., "*/fri:t/*") followed by a prime word. The subjects tended to write down "freat" and "freet" when the prime words were "cheat" and "greet," respectively. Delattre et al. (2006) manipulated the POC regularity and lexical frequency of target words in a writing to dictation task with healthy French speakers, and observed the positive main effects of these two variables in response latencies. Furthermore, there was a larger regularity effect for low-frequency than for high-frequency words, indicating the interaction effects between these two variables. Using the multiple regression study approach, Bonin and Méot (2002) asked healthy participants to spell 150 monosyllabic French words to dictation, and further analyses of the data (Bonin et al., 2004) showed that the spelling to dictation latencies could be significantly predicted by POC transparency (e.g., phonology–orthography consistency of the rime units), and certain lexical factors (e.g., rated AoA, cumulative frequency), but not by any factors indicating semantic FAM and/or richness (i.e., conceptual IMG, CON, and FAM; see similar results obtained with primary school French children: Lété, Peereman, & Fayol, 2008).

It is interesting that, although the effects of lexical variables (e.g., frequency, AoA, lexical priming) and POC variables have been consistently reported in writing studies with healthy adults, the effects of semantic variables have not. Although Bonin and Méot (2002) reported such effects, the results did not hold in the subsequent analyses with an improved method (Bonin et al., 2004). This is in contrast with the writing patterns of deep dysgraphic patients, where semantic errors are prevalent (e.g., Bub & Kertesz, 1982; Duhamel & Poncet, 1986; Majerus et al., 2001; Miceli et al., 1997; Tainturier & Caramazza, 1996; Valdois et al., 1995; Weekes & Raman, 2008). Bonin et al. (2004) discussed this difference and attributed the lack of semantic effects to language specificity. They proposed that the extent of involvement of semantic representation in spelling to dictation is modulated by the degree of POC transparency of a particular language. For alphabetic languages with high POC transparency (e.g., Italian, Turkish), their normal spelling to dictation could be accomplished through the POC route coupled with some lexical information and the semantic involvement could be minimal. This proposal explains the semantic effects in deep dysgraphic patients readily: in those cases the POC (and lexical) routes are compromised and cannot generate outputs for writing accurately. By extension, in languages with more opaque POC correspondence (e.g., with high degree of homophony), the involvement of the semantic properties might become more apparent. Indeed, Bonin et al. (2004) contemplated that "semantics may play a role in spelling when these representations are particularly useful to disambiguate the spelling of heterographic homophones (such as rain, rein, and reign)" (p. 471).

Chinese is an ideal language to examine this prediction. Chinese is a logographic script with no POC correspondence on the segmental level and with a high degree of homophony. The basic writing units in Chinese are characters, most of which are either free morphemes (i.e., words when standing alone, about 65%) or bound morphemes (Sun, Huang, Sun, Li, & Xing, 1997). Each character (e.g., 疤, scar) corresponds to a syllable in sound (*/ba1*[1]*/*). There is no internal "phoneme" in the syllable that corresponds to any target "grapheme." Furthermore, a syllable maps to 15 homophonic characters (i.e., characters that have identical pronunciations but different meanings) on average (Standards Press of China, 1994). An extreme case is the syllable */yi4/*, which corresponds to 205 homophonic characters (e.g., 意-meaning, 议-discuss, 亿-100 million, 易-easy, 艺-art, 译-translate) (Standards Press of China, 1994). Shen and Bear (2000) have shown that Chinese children made a small proportion of semantic errors (around 5%) in addition to phonological and orthographic ones in writing essays and writing to dictation with contextual constraints (for similar results in English-speaking children, see also Sadoski, Willson, Holcomb, & Boulware-Gooden, 2004). However, such results do not conclusively support Bonin et al.'s prediction because (a) their tasks emphasized semantic processing heavily and (b) the cognitive mechanism used during development may be different from those used by matured systems. The current study explores the mechanisms underlying normal writing to dictation in healthy Chinese speakers by examining the effects of different variables in this process. The primary aim is to test the prediction by Bonin and colleagues (2004) that semantic effects would be more easily detected in such a language. A further question is whether the potential semantic route interacts with the POC route in Chinese writing to dictation, as well as the potential interaction between lexical and POC routes.

Writing theories of Chinese similar to those of alphabetic languages have been proposed. There are three possible routes in Chinese spelling system: semantically mediated lexical, direct lexical, POC ones. The two lexical routes (i.e., a direct lexical route and a semantically mediated lexical route) have been incorporated in Chinese writing model (for a review, see Weekes, Yin, Su, & Chen, 2006). By contrast, it remains controversial whether Chinese writing relies on the POC route. We here assume that such a nonlexical POC route exists in Chinese writing system for the following reasons. First, the POC route works on the rule-based mapping probability from auditory units to written form units. This mapping probability is derived based on the statistical distributions in the language. In this sense, the working principle of the POC route in Chinese is analogous to that of the POC route in alphabetic spelling system (see below for further discussion on this point), and both are not lexical. Second, mapping mechanisms based on the statistical probabilities are domain-general in the human's cognitive processes. An abundance of evidence has revealed that the mapping processing on basis of statistical learning could occur on both linguistic and nonlinguistic stimuli (e.g., Fiser & Aslin, 2002; Kirkham, Slemmer, & Johnson, 2002; Saffran, Aslin, & Newport, 1996; Saffran, Pollak, Seibel, & Shkolnik, 2007). Moreover, Lee and colleagues (e.g., Lee, Huang, Kuo, Tsai, & Tzeng, 2010; Lee, Tsai, Chiu, Hung,

& Tzeng, 2006; Lee et al., 2004) also found that there is a mapping procedure (POC) in Chinese reading, which is treated as the counterpart of writing. Hence, it is possible that Chinese spelling system has a similar statistical mapping procedure, that is, the POC route.

Note that clearly in Chinese the phoneme–grapheme POC procedure in alphabetic scripts does not directly apply. Rather, we adapt its notion as a rule-based mapping system between sound units and written form units. Although in alphabetic scripts the functional units for POC are phonemes and graphemes, in Chinese they are syllables and characters. We acknowledge that these two types of mappings in different scripts should have different mechanisms. Here we simply attempt to capture a possible way of sound–written form conversion on the basis of mapping probability. We assume that the syllable–character correspondence "rules" are derived based on the statistical distributions in the language, measured by the frequency of the specific mapping (estimated by the frequency of the character) in proportion to the syllable frequency and the consistency of the mapping (estimated by the number of characters corresponding to the same syllable). Furthermore, we have assumed that this POC route can function purely on the basis of mapping probability and could be separated from the meaning of the particular character (morphemes) and/or other lexical characteristics of the character (morpheme). In this case, we examine whether the writing to dictation process in Chinese is influenced by the mapping probability between sounds and visual forms, and this particular mechanism interacts with the *semantic* route, in which the meaning of the characters/words plays a role (see a similar discussion for the OPC route in reading in Bi, Han, Shu, & Weekes, 2007).

In the present study, we employed a writing to syllable task that exhausted all Chinese syllables with frequency counts higher than 0, where Chinese participants were asked to write down the first character (morpheme) that came to mind upon hearing a syllable. We used a multiple regression analysis approach to evaluate the effects of relevant psycholinguistic variables (i.e., lexical, semantic, POC ones, and the interactions) in predicting the production likelihood of target. In addition to this main analysis for our primary research aim, we also carried out two sets of additional analyses. One is to explore the nature of the observed semantic effects in the main analyses in Chinese writing. In particular, we tested in greater detail whether the semantic results are related to the language or to the specific stimuli set. A further analysis is to inspect any potential role of the feedback POC mechanism—the orthography to phonology mapping—at both whole character and subcharacter levels (see Method and Result sections for more details). Finally, a supplementary outcome of the current study is a database of responses in writing Chinese syllables to dictation available for future study.

## METHOD

### *Participants*

Thirty college students from Beijing Normal University took part in the present study (15 females, 15 males; mean age = 21 years, age range = 19–24 years). All were native Mandarin Chinese speakers with normal or corrected to normal vision and normal hearing. Each subject received the full set of stimuli. They gave informed consent, understood the rights to withdraw any time from the experiment, and received a payment for participation.

### *Material*

To cover as many syllables as possible, we referred to one of the largest character database. It contains 20,902 characters (UCS Chinese character database; Standards Press of China, 1994). The database includes 1,407 syllables (tones are considered as distinctive features), among which 1,190 have syllable frequencies (the sum of frequencies of characters corresponding to the syllable) higher than zero. A female native speaker of Mandarin Chinese recorded the audio files of the 1,190 syllables in a soundproof recording studio. We chose a high-frequency character corresponding to each syllable. The characters were presented in a random order for her to read aloud and the digital recording was later transformed into individual audio files. Each file corresponded to a syllable. The files were evenly assigned into five blocks for different testing sessions in a pseudorandom fashion. No syllables with identical segments (onset and rime) occurred in a same block.

### *Procedure*

Each subject completed the whole experiment in five sessions on five different days, with each session lasting for about 30 min. The order of the five stimuli blocks was counterbalanced across subjects in a Latin-square manner. The subjects were individually tested in a quiet room. The syllabic audio files were presented over a headphone using the DMDX software (Foster & Foster, 2003) on a Dell laptop (Model Inspiron 6000). The subjects self-paced the stimuli presentation with the space bar. They were instructed to write down on a piece of paper the first character that came to mind upon hearing a syllable. Each syllable was presented in isolation rather than in sentence context because we did not want to provide any semantic constraint. The pen was attached to the digital tablet (PTZ-631W). In case of failure to generate any target character, he or she was asked to draw a circle or write down an associated character, which can form a Chinese two-character word with the potential character that corresponds to the target syllable (an example in English would be to write "-fly" for the target */butter/*).

### *Description of the variables*

We treated each character (response) that was correctly produced for a syllable in the experiment as an item in the regression analyses. The dependent variable was the probability of a character being generated for a given syllable target across the subject sample (i.e., generation probability [GP]; see Table 1), which was the number of subjects producing this character divided by the total number of subjects (i.e., 30). Take the syllable */wo4/* for example: 27 subjects generated correct responses (13 握, */wo4/*, grasp; 12 卧, */wo4/*, lie down; 2 沃, */wo4/*, fertile), one wrote a character that was phonologically similar to the target (窝, */wo1/*, nest), and two failed to supply any response. The three correct character responses

Table 1. *Mean, standard deviation, and range values and correlation matrix (Pearson r) for the dependent measure (GP) and the predictor variables*

| Predictors | Mean | *SD* | Range | GP | FT | rAoA | oAoA | WF | IMG | CON | FAM | HF |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Lexical level | | | | | | | | | | | | |
| FT | 0 | 0.60 | −4.36–0.68 | −0.07** | | | | | | | | |
| rAoA | 4.14 | 0.80 | 1.86–6.59 | −0.45** | 0.10** | | | | | | | |
| oAoA | 4.75 | 3.64 | 1.00–22.00 | −0.30** | 0.05* | 0.65** | | | | | | |
| WF | 1.42 | 0.76 | 0–4.25 | 0.47** | 0.01 | −0.65** | −0.51** | | | | | |
| Semantic level | | | | | | | | | | | | |
| IMG | 5.47 | 0.90 | 1.97–7.00 | 0.28** | −0.23** | −0.45** | −0.25** | 0.26** | | | | |
| CON | 5.09 | 1.03 | 2.20–7.00 | 0.12** | −0.26** | −0.30** | −0.13** | 0.11** | 0.81** | | | |
| FAM | 6.15 | 0.61 | 2.91–7.00 | 0.31** | −0.09** | −0.60** | −0.40** | 0.44** | 0.61** | 0.49** | | |
| POC level | | | | | | | | | | | | |
| HF | 0.78 | 0.19 | 0.14–1.00 | 0.66** | 0.08** | −0.42** | −0.34** | 0.52** | 0.07** | −0.07** | 0.27** | |
| NH | 6.38 | 5.24 | 1–40 | −0.44** | 0.09** | 0.02 | −0.05* | −0.08** | −0.17** | −0.16** | −0.05 | −0.39** |

*Note:* GP, generation probability; FT, frequency trajectory; rAoA, rated age of acquisition; oAoA, objective AoA; WF, word frequency; IMG, imageability; CON, concreteness; FAM, familiarity; HF, homophone frequency; POC, phonology-to-orthography conversion; NH, number of homophones.

*$p < .05$. **$p < .01$.

(握, 卧, and 沃) entered further analyses where their GP values were 43% (13/30), 40% (12/30), and 7% (2/30), respectively.

For the predictors, we followed those in Bonin et al. (2004) that were potentially relevant to the spelling to dictation process outlined in the Introduction: lexical ones, semantic ones, and ones that reflected the POC correspondence (see Table 1). Note that some variables (e.g., AoA) could be considered either as lexical (e.g., Gerhand & Barry, 1998; Lewis, Gerhand, & Ellis, 2001), or as semantic (e.g., Brysbaert, van Wijnendaele, & de Deyne, 2000; Steyvers & Tenenbaum, 2005) depending on the specific theories, we here follow a conventional way of variable categorization. Below the selection and calculation method of each variable is presented in detail.

*Lexical level variables.*

FREQUENCY TRAJECTORY (FT). FT refers to how experience with a word is distributed over time (Zevin & Seidenberg, 2002). It has been suggested that this reflects the learning effect better than the conventional AoA measures (see Bonin et al., 2004; Bonin, Méot, Mermillod, Ferrand, & Barry, 2009). Following Bonin et al. (2004; see also Cuetos & Barbon, 2006; Pérez, 2007), we defined FT for a character as the difference between the $z$ scores of its frequency in the adult corpus (Sun et al., 1997) and that in the children corpus. The children corpus includes all characters in the standard textbooks that are used nationally in elementary schools (G1–G6; People's Education Press, 2006). To avoid confounding effects from different size datasets in counting the two $z$ scores, we calculated their $z$ values on basis of one same dataset, that is, the 1,705 character set consisted of items entering our regression analyses. A positive FT value for a character meant that the trajectory of the frequency increased during development, that is, the character is of low frequency in the early period and of high in the later period. A negative FT value meant the reverse.

RATED AoA. The rated AoA values were taken from the Chinese Single-Character Word Database (CSWD; Liu, Shu, & Li, 2007). In this database, Liu et al. (2007) asked adults to estimate the age at which they had learned the word in either their spoken or written form, following the method used in Barca, Burani, and Arduino (2002). This database contains measures of 15 variables for 2423 Chinese monosyllabic words.

OBJECTIVE AoA. Taken from Liu et al. (2007), these values show the time when a character was first introduced in the standard school textbooks (People's Education Press, 2006).

WF. The log value of the frequency count of a character as a monosyllabic word in Sun et al. (1997).

*Semantic level variables.* We took the values of three semantic variables from the CSWD norm (Liu et al., 2007): IMG, CON, and FAM. They were based on subjective ratings on a 7-point scale adapted from earlier studies (e.g., Barca et al., 2002). For IMG ratings, subjects were told that IMG is the degree to which the word arouses a mental image (i.e., a mental picture, sound, or other sensory experience). Their task was to rate each word for its degree of IMG (1 = *hardly imageable*, 7 = *highly imageable*). For CON ratings, they were told that the words that refer to objects, living beings, actions, and materials that can be experienced by the senses can be considered as "concrete" and the words that refer to concepts that cannot be experienced by the senses can be considered as "abstract." Their task was to rate each word for its degree of CON/abstractness (1 = *highly abstract*, 7 = *highly concrete words*). For FAM ratings, they were required to rate how well they thought each word was known by students like them (1 = *not known*, 7 = *very well known*).

*POC level variables.* Although there is no POC on segmental levels between syllables and characters in Chinese, we can nonetheless measure the likelihood that a certain written form will correspond to a given sound (syllable) without considering any lexical semantic characteristics of the character (see further considerations for this point in Introduction). The following measures attempt to capture such PO mapping probability.

HOMOPHONE FREQUENCY (HF). This measure parallels the HF measure in English, and was the ratio of the log frequency of a specific character to the log frequency of its syllable (sum frequency of all characters corresponding to this syllable). Both syllable frequency and character frequency were derived from Sun et al. (1997). Take the syllable */ba4/*, for example. There are four characters corresponding to it in the database. The four characters and their character frequencies were 爸: 776; 罢: 95; 坝: 49; and 霸: 30. The syllable frequency of this syllable was then 950 (776 + 95 + 49 + 30). The HF values of the characters were lg (776)/lg(950), lg (95)/lg(950), lg (49)/lg(950), and lg (30)/lg(950), respectively.

NUMBER OF HOMOPHONES (NH). The number of characters (i.e., homophones) corresponding to a syllable was also considered because it might reflect the degree of competition for a character response in addition to HF. In particular, Bonin et al. (2004) speculated that the importance of the semantic variables is modulated by the degree of homophony (see Introduction). The NH for a given character in our analysis was calculated on the basis of the Sun et al. (1997) database. Take the syllable */ba4/*, for example. It corresponds to four characters in the database and the NH for each of the characters was 4.

*Interactions.* To elucidate the potential modulatory effects of between the lexical semantic route and POC route in writing process, we also included the interactions between them as predictors in the regression analyses. The detailed methods for this part are presented in the Results Section.

To fully understand the contribution of different kinds of variables on GP, we carried out multiple regression analyses in a number of ways evaluating the effects of the cognitive routes. The regression methods included the simultaneous Enter method and the two-step hierarchical method. Identical analyses were conducted for two different set of variables: individual and the principal component analysis (PCA) ones. Individual variables were selected as the best representative variable at each level and PCA ones were extracted from each level on the basis of PCA procedure.

## RESULTS AND DISCUSSION

All 30 subjects completed the full set of stimuli (writing to dictation of 1,190 syllables). Any character responses that correspond to the target syllable in the UCS Chinese Character Database (Standards Press of China, 1994) were considered correct. Erroneous responses included (a) no responses, (b) characters associated with a potential character that corresponds to the target syllable, (c) characters that were phonologically related to the target syllable (e.g., */ding4/* -> 令, */ling;4/*, command), and (d) writing errors (e.g., stroke deletion or insertion). All the subjects correctly wrote down 29,979 responses in total, corresponding to 2,807 characters. A same character might be written by several subjects. For example, the character "爸" (dad) was generated by 21 subjects for the syllable "*/ba4/*," and "马" (horse) by all 30 subjects for "*/ma3/*." The mean accuracy across subjects was 84% (*SD* = 4%, range = 70%–91%). On average, every target syllable yielded 2.4 (2,807/1,190) correct character responses across subjects. The syllable */ji4/* had very diverse responses, yielding 11 correct character responses: 记 (record), 继 (subsequent), 计 (calculate), 既 (therefore), 纪 (rule), 忌 (envy), 季 (season), 寂 (lonely), 冀 (hope), 系 (tie) and 迹 (trace). About 3% (34/1,190) syllables did not generate any correct responses due to their extremely low frequency. For instance, the syllable */cuo3/* with a syllable frequency of 1 was miswritten by all subjects; its erroneous responses included 18 no responses, 8 phonologically related characters, 3 associated characters, and 1 writing error.

### *Normative data*

We first established a norm of writing to dictation of Chinese syllables based on the complete list of our data (available at http://psychbrain.bnu.edu.cn/home/yanchaobi/writing-to-dictation.xls). For each syllable, this norm shows all the correct character responses produced by the 30 subjects, along with the number of subjects that produced them. Take syllable */huo4/*, for example; the norm file shows that it was produced correctly by 29 subjects. The 29 correct responses corresponded to five characters. The character responses and their corresponding generation frequency (number of subjects that produced it) were 或 (19), 祸 (4), 货 (3), 霍 (2), and 获 (1). This database could potentially be used as a reference for the understanding of phonology–orthography mapping in Chinese writing beyond the target issues in the current study.

### *Effects of lexical, semantic, and POC attributes*

Out of the 2,807 correctly produced characters, 1,705 are monosyllabic words that were included in the CSWD norm (Liu et al., 2007) and the other databases we adopted. We carried out our analyses based on these 1,705 characters

Table 2. *Results of the regression analyses (simultaneous entering method) for individual variables and PCA variables*

| Predictors | *B* | *SE* | Beta | *t* | *p* |
|---|---|---|---|---|---|
| Individual Variables | | | | | |
| Lexical FT | −0.04 | 0.01 | −0.07 | −3.96 | <.0001 |
| Semantic IMG | 0.08 | 0.01 | 0.23 | 12.65 | <.0001 |
| POC HF | 1.12 | 0.03 | 0.65 | 37.25 | <.0001 |
| PCA Variables | | | | | |
| Lexical PCA | −0.09 | 0.01 | −0.27 | −13.73 | <.0001 |
| Semantic PCA | 0.26 | 0.01 | 0.08 | 4.152 | <.0001 |
| POC PCA | 0.19 | 0.01 | 0.57 | 31.90 | <.0001 |

*Note:* PCA, principal component analysis; FT, frequency trajectory; IMG, imageability; POC, phonology-to-orthography conversion; HF, homophone frequency.

(monosyllabic words) because they had the most complete linguistic variable values across the databases we used. Following the description in the Method section, the dependent variable in our regression analyses was the GP of each character. The predictive variables were nine relevant attributes of the characters in three levels of writing to dictation processing: lexical, semantic and POC ones.

*Correlation between the variables.* Table 1 displays the mean values of the variables and the Pearson's correlation matrix. The dependent variable (GP) had significant correlations with all predictive variables ($p$s $< .01$). Independent variables within the same level generally correlated with each other highly ($p$s $< .01$), except for the correction between FT and WF ($r = .01$, $p > .05$), which was not surprising (see Bonin et al., 2004; Zevin & Seidenberg, 2002). Words with same WF might experience different distribution of WF over time (i.e., FT). Some words might change from being high frequency in early period to being low frequency in late period, other words be quite the reverse, or equivalent in different periods. Therefore, it is possible that there is a low correlation between FT and WF.

*Effects of individual routes.* To examine the prediction of Bonin et al. (2004) that semantic variables are more likely to influence writing to dictation in languages with deep orthography, we evaluated the contribution of the three levels (lexical, semantic, and POC) to the dependent variable (GP) using two sets of variables: individual and PCA ones (see Table 2). For individual variables, we selected the most representative variable for each route according to the convention of the literature. The selected variables were FT for lexical level (e.g., Bonin et al., 2004, 2009), imageability for semantic level (e.g., Bird, Howard, & Franklin, 2003), and homophone frequency for nonlexical level (e.g., Rapp et al., 2002). For PCA variables, we employed the PCA procedure (see Baayen, Feldman, & Schreuder, 2006) to extract one single composite factor for each route from all corresponding variables described in the method section. The PCA technique is a mathematical procedure that transforms a number of possibly correlated variables into a smaller number of uncorrelated variables called principal components. To obtain the PCA factor at one level (e.g., semantic level), we entered all variables (IMG, CON, and FAM) at this level into PCA program. We used the subcommands for varimax rotation, a plot of the eigenvalues, and a principal components extraction. Three PCA factors were therefore obtained for the three routes: lexical, semantic, and POC ones. These PCA factors captured 55%, 76%, and 70% of the variances for the three levels, respectively. The correlations of the PCA factors were all lower than .70 ($R_{\text{lexical PCA-semantic PCA}} = -.43$; $R_{\text{lexical PCA-POC PCA}} = -.31$; $R_{\text{semantic PCA-POC PCA}} = .15$). As pairwise correlation lower than .70 has been proposed as the rule of thumb generally adopted for regression studies (Baayen et al., 2006), our regression analyses on PCA variables would be effective.

We then carried out multiple regression analyses using both "Enter" and hierarchical regression methods with identical procedures across these two variable sets. We first used the simultaneously "Enter" method to introduce the variables of the three levels as predictors and GP as the dependent variable into the regression equation. All predictors in the individual variable set had significant explanatory power for GP ($p$s $< .0001$; Table 2). The earlier acquired, the more imageable, and the more homophones it has, the more likely the character was produced. All predictors of PCA variable set had significant explanatory power for GP ($p$s $< .0001$; Table 2). These results confirmed that all three routes are involved in deriving the writing to spelling output in Chinese.

To further examine the unique contribution of each route by ruling out the effects of the other two, we conducted the two-step hierarchical regression analyses. This was done by entering the variables of two routes into the regression model in the first step, and then the variable of the route of interest in the second step. The value of the $\Delta R^2$ of regression models from Step 1 to Step 2 reflects the contribution of the variable entered in the second step. The results using this method are shown in Table 3. We found that the $\Delta R^2$s for all three levels examined this way for the two sets of variables (individual and PCA) were both statistically significant ($p$s $< .0001$), further confirming that the lexical, semantic, and POC factors all influence the writing to dictation outcomes for Chinese syllables.

The above results also showed that different kinds of variables had different degrees of contribution on GP (see Tables 2 and 3). In general, both the semantic and lexical effects on GP were weaker than the POC effect. Worth mentioning is that although some lexical and semantic variables had low beta or $\Delta R^2$ values (e.g., FT $= -0.07$, semantic PCA $= 0.08$), the effects were nonetheless rather stable and significant ($p$s $< .0001$ across several regressing analyses).

*Effects of the interlevel interaction.* The previous studies have observed the interaction effects between lexical and POC factors in the spelling process of healthy adults in alphabetic scripts (e.g., Barry & Seymour, 1988; Bosse et al., 2003; Campbell, 1983; Delattre et al., 2006) and in the development processing of writing in Chinese (Shen & Bear, 2000). The interaction effects between semantic

Table 3. *Results of the second-step $R^2$ changes (two-step hierarchical method) using the individual PCA variable as the second-step predictors*

| Predictors Entered in Second Step | $\triangle R^2$ | $\triangle F$ | $p$ |
|---|---|---|---|
| Individual Variables | | | |
| Lexical FT | 0.005 | 15.650 | <.0001 |
| Semantic IMG | 0.048 | 159.99 | <.0001 |
| POC HF | 0.413 | 1387.31 | <.0001 |
| PCA Variables | | | |
| Lexical PCA | 0.053 | 188.49 | <.0001 |
| Semantic PCA | 0.005 | 17.24 | <.0001 |
| POC PCA | 0.288 | 1017.53 | <.0001 |

*Note:* PCA, principal component analysis; FT, frequency trajectory; IMG, imageability; POC, phonology-to-orthography conversion; HF, homophone frequency.

factors and POC factors in dysgraphic writing have also been demonstrated (e.g., Hillis & Caramazza, 1991, 1995; Ward et al., 2000).

Here we examined the potential modulations between semantic and nonlexical routes, and between lexical and nonlexical routes in predicting the GP values, adopting the two-step hierarchical regression analysis. In the first step of the analysis, we entered all three level factors into the regression model, and then in the second step, we examined whether the inclusion of either of the two interaction indexes (Semantic × POC; Lexical × POC) improved the explanatory power significantly in two separate analysis. For the two sets of variables, the two interaction indexes induced significant $\Delta R^2$ ($p$s $< .0001$; Table 4), suggesting that the interaction between different processing levels has effects beyond and above the effects of individual processing levels in predicating the outcome in Chinese syllable writing to dictation. Specifically, the less efficient POC procedure is, the more heavily both semantic and lexical routes are relied on. This provides new evidence of the existence of interaction between lexical semantic and nonlexical routes in Chinese writing.

### *The nature of the semantic involvement in Chinese writing*

In the above analysis we have established that semantic factors play a significant role in normal Chinese writing to dictation. Here we examine whether the semantic involvement in writing Chinese is related to the language or to the specific stimuli set that was chosen. According to Bonin et al.'s (2004) speculation, the stimuli play a critical role in how much semantic properties affect the writing to dictation process. Nonhomophonic stimuli, as in Bonin et al. (2004), will not exhibit a semantic effect in writing.

Table 4. *Results of the second-step $R^2$ changes (two-step hierarchical method) using the interaction PCA variable as the second-step predictors*

| Predictors Entered in Second Step | $\triangle R^2$ | $\triangle F$ | $p$ |
|---|---|---|---|
| Individual Variables | | | |
| Lexical FT × POC HF | 0.008 | 26.76 | <.0001 |
| Semantic IMG × POC HF | 0.008 | 28.41 | <.0001 |
| PCA Variables | | | |
| Lexical PCA × POC PCA | 0.026 | 102.93 | <.0001 |
| Semantic PCA × POC PCA | 0.008 | 28.43 | <.0001 |

*Note:* PCA, principal component analysis; FT, frequency trajectory; POC, phonology-to-orthography conversion; HF, homophone frequency; IMG, imageability.

To test this possibility, we ran the regression using the enter method on a subset of 190 characters in our stimuli set that had no homophones (NH = 1). The predictors were the semantic and lexical PCA factors, and the dependent variable was GP. The results showed that the semantic factor had no predictive power for GP ($p = .65$) while the lexical one had a significant effect ($p < .0001$). This indicates that the semantic involvement in writing to dictation is stimuli related, verifying the hypothesis proposed by Bonin et al. (2004) that semantically mediated writing is a dominant strategy only in cases with many homophones instead of a language-specific feature.

### *Effects of the OPC (feedback) procedure in Chinese writing*

Until now we have focused on the phonology to orthography correspondence when discussing nonlexical conversion procedures of writing to dictation following the convention. Here we further consider the potential effect of feedback, that is, orthography to phonology correspondence (OPC). In Chinese, there are potentially two manners in which such a feedback procedure might occur: at the whole-character level and at the subcharacter (i.e., radical) level. The feedback effect at the whole-character level applies in the case of homography. A homographic character (e.g., 会) could have two or more pronunciations (e.g., */hui4/*, able; */kuai4/*, accounting). Characters with one, two, and three pronunciations account for 68%, 23%, and 10%, respectively, in the UCS Chinese Character Database (Standards Press of China, 1994), respectively. It is possible that, compared with characters with a single pronunciation, those with multiple pronunciations are more likely to be selected for writing because they benefit from the activation from multiple pronunciations through feedback loops. The feedback mechanisms at the subcharacter level mainly come from so-called composite characters. Over 80% of modern characters are phonetic composite characters (e.g.,"妈", */ma1/*, mother) comprising two components: a semantic radical (e.g. "女", female) that provides clue about the meaning of the character, and a phonetic radical ("马", */ma3/*) that gives information about the character's pronunciation. To characterize how reliably the pronunciation of a composite character can be predicted by the sound of its phonetic radical, two measures are commonly used: regularity and consistency. Regularity refers to the degree to which a whole character shares pronunciation with its radical. Consistency refers to whether the pronunciation of a character agrees with those of its orthographic neighbors containing the same phonetic radical (see Law, 2004; Law, Wong, & Kong, 2006; Law, Yeung, Wong, & Chiu, 2005; Shu, Chen, Anderson, Wu, & Xuan, 2003). Some authors have demonstrated that information of semantic and phonetic radicals involves in Chinese writing in dysgraphic population. For instance, Law et al. (2005) observed that their patient (NMY) made many writing errors on semantic radicals, including semantic radicals' institution, insertion, and deletion. Moreover, the substituting and inserted semantic radicals were semantically related with the target words, indicating that semantic radicals could be decomposed from the whole characters and are directly connected with semantic features.

To explore the existence of a feedback OPC effect at whole-character level for writing to dictation in Chinese, we conducted additional regression analyses on our data set. Among the 1,705 characters in the actual corpus, characters with one, two, and three pronunciations were 56%, 27%, and 18%, respectively. For each character, the dependent variable was GP, and the predictor was the number of pronunciations in UCS Chinese Character Database (Standards Press of China, 1994). When only the number of pronunciation predictor was entered in the regression model, it showed a significant effect ($t = 3.10$, $p = .002$), which diminished ($p$s $> .05$) once either the other variables (FT, IMG, and HF) or the PCA variables (lexical, semantic, and nonlexical PCAs) were included. That is, we did not observe any significant effect of OPC at the whole-character level.

We then considered the OPC mechanisms at the subcharacter level. Three indexes are commonly used to capture such a procedure: (a) the phonetic radical transparency: the extent to which the pronunciation of a whole character is denoted by the phonetic radical; (b) the phonetic radical consistency: the degree of consistency of the pronunciations of the family of characters sharing the same phonetic component; and (c) the semantic radical transparency: the extent to which the meaning of a whole character can be derived from the semantic radical. To examine whether these indexes play a role in writing to dictation of Chinese, specific regression analyses were conducted on the 920 composite characters in our dataset. These characters were listed in the elementary Chinese character corpus (see Shu & Wu, 2006; Xing, Shu, & Li, 2004), which provides values of the three subcharacter OPC indexes as follows. The phonetic radical transparency was labeled according to the relationship of pronunciations between the whole character and its phonetic radical: (a) same by onset, rime, and tone; (b) same by onset and rime; (c) same by rime; (d) same by onset; or (e) completely different. The phonetic radical consistency was labeled as follows: (a) there is only one character with this phonetic radical in the database; (b) all characters with the same phonetic radical are pronounced the same; or (c) characters with the phonetic radical are pronounced differently. Similarly, the semantic radical transparency was coded from 1 to 8 according to the similarity between the meaning of the radical and the whole character. The potential effects of each of these three variables were examined using regression in the same way as the number of pronunciations. The phonetic radical transparency consistently showed significant predictive power for GP in all analyses ($p$s $< .001$), whereas the phonetic radical consistency and the semantic radical transparency did not exhibit significant effects in any analysis ($p$s $> .10$).

## GENERAL DISCUSSION

We investigated the cognitive mechanism for writing Chinese characters from dictation using a large-sample multiple regression method with healthy adult subjects. The question was when a syllable corresponds to multiple characters (homophones), what attributes of a particular character affect the probability of it being selected for output. We observed significant predicative effects of all three types of variables—lexical, semantic, and POC—as well as interactions between the POC variables and the other two. The same results were obtained both when we used individual variables as predictors and when we used factor clusters extracted by means of PCA. Moreover, when a syllable only corresponds to one character, the semantic effect was not visible any more. We also observed the effect of an orthography–phonology–conversion transparency variable—the transparency of phonetic radical—in writing to dictation of Chinese, indicating the existence of a feedback mechanism.

The implications of our results are multifold. The results on the writing to dictation process in healthy adults provide converging evidence for the writing theories discussed in the Introduction. The effects of semantic variables and POC variables and their interaction have been documented in dysgraphic writing, but in the spelling process of healthy subjects the semantic effects were not established. To our knowledge, the current study reports the first piece of evidence showing that the normal adult writing to dictation system also involves semantic processing. Our results are consistent with Bonin et al.'s (2004) proposal that although the semantic effect is hardly visible in relatively transparent orthographies (e.g., Italian, Turkish), it should be more easily detectable in opaque languages (e.g., Chinese). These authors contended that the semantic representation would participate in spelling to dictation especially when the POC procedure is inefficient in disambiguating the selection of alternative orthographic outputs—due to opaque POC and/or existence of many POC legitimate candidates. Following this proposal, the result discrepancies between the current study and Bonin et al.'s (2004) study with similar approaches could be readily attributed to cross-linguistic differences between languages with deep versus medium to transparent orthographies. This could also explain the differences obtained across subject groups speaking the same type of languages, such as the existence of semantic effects in writing in alphabetic scripts by dysgraphic patients (i.e., deep dysgraphia in French, English, Italian, and Turkish; e.g., Bub & Kertesz, 1982; Duhamel & Poncet, 1986; Majerus et al., 2001; Miceli et al., 1997; Valdois et al. 1995; Weekes & Raman, 2008) but not in healthy subjects (e.g., Bonin et al., 2004). There was positive evidence that deep dysgraphic patients suffer from deficits for POC and lexical information (e.g., Hillis et al., 1999). Our results showed directly that the semantic involvement in writing Chinese is a characteristic for the specific stimuli set with multiple homophones. When the stimuli had no homophones, semantic effects disappeared. This is consistent with Bonin et al.'s (2004) speculation that the semantics route is used to disambiguate the spelling of heterographic homophones.

Another important finding in our study is that the interactions between semantic factors and POC factors showed a significant contribution to predicting how likely a particular character is produced given a syllable target. In alphabetic scripts, only interaction between lexical and POC variables have been documented in healthy adult writing. For example, Delattre et al. (2006) observed that healthy French speakers in writing to dictation task showed significant interaction between POC regularity and lexical frequency of target words, that is, a larger regularity effect for low-frequency than for high-frequency words (for parallel results in neuropsychological studies, see also Rapp et al., 2002). In our study, after eliminating the effects of individual factors, we still observed significant interaction effects between semantic factors and POC factors (in addition to between lexical factors and POC factors) in predicting the character GP. Again, this is in accordance with the dysgraphic performances in patients with brain damage in alphabetic writing systems, where interaction between semantic and POC factors were observed (e.g., RCM in Hillis et al., 1999). In summary, the similarity in result patterns between healthy Chinese speakers in our study and the dysgraphic patients in the literature of writing to dictation in alphabetic scripts might be because in both circumstances the POC route is not effective in constraining writing outcomes. Specifically, in the theoretical framework of Chinese writing, for a given auditory stimuli, the POC route activates the target and the phonological-related items (including homophones), whereas the semantic–lexical route actives the target and the semantic-related items. The semantic–lexical route may not have an effect when the POC route by itself could efficiently select the output candidate (such as for the items without homophones). However, the semantic–lexical route has to participant in written selection when the POC route fails to generate the final representation (such as for the items with many homophone or when the POC route is damaged).

We also observed that an orthographic variable—the phonetic radical transparency—influences the processing of writing to dictation in Chinese, indicating that this feedback mapping contributes to healthy Chinese writing to dictation system. This piece of result provides new evidence complementing Law et al.'s (2005) finding that semantic and phonetic radicals of composite characters could be decomposed from the character and participant in Chinese written production. It is also consistent with the findings in alphabetic languages showing the effect of certain feedback factors (e.g., phoneme to morpheme transparency) on writing (e.g., Lété et al., 2008; Ziegler & Ferrand, 1998).

One caveat to consider is that not all written responses were included in the data analyses. We only analyzed 60% (1,705 out of 2,807 correctly produced characters) of all responses because only these response characters had the most complete linguistic variable values across the databases employed. This might be a bias selection as the unanalyzed characters might have lower frequency, be less imageable/concrete, and so forth. If the whole data set had been entered into the regression equations, the effects observed in the present analyses might be less well defined. Further studies are warranted.

To conclude, we observed that in addition to lexical variables and POC probabilities, semantic variables influence how likely a character is produced out of a homophone family in syllable writing to dictation by healthy Chinese speakers. Furthermore, the semantic and POC routes modulate each other. Such results confirm the prediction by Bonin et al. (2004) that the extent of the semantic involvement is influenced by POC transparency, and corroborate the writing theories developed on the basis of cognitive and neuropsychological studies across different subject groups. The results also suggest that writing mechanisms could be stimulus specific rather than language specific.

## ACKNOWLEDGMENTS

This research was funded by grants from the Changjiang Scholars and Innovative Research Team in University (IRT0710), National Natural Science Foundation of China (30770715, 30700224), National Social Science Foundation of China (07CYY009), and Beijing Natural Science Foundation (7082051). We thank Mingxuan Tan, Manyun Zhang, and Yue Wang for assistance in data collection and Don Cantor for proofreading the paper. The first and second authors contributed equally to this work.

## NOTE

1. The digit represents tone of the preceding syllable. There are four tones in Mandarin Chinese. Only syllables with identical, onset, rime, and tone are considered as homophones.

## REFERENCES

Baayen, R. H., Feldman, L. F., & Schreuder, R. (2006). Morphological influences on the recognition of monosyllabic monomorphemic words. *Journal of Memory and Language, 53*, 496–512.

Barca, L., Burani, C., & Arduino, L. S. (2002). Word naming times and psycholinguistic norms for Italian nouns. *Behavior Research Methods, Instruments, and Computers, 34*, 424–434.

Barry, C., & Seymour, P. H. K. (1988). Lexical priming and sound-to-spelling contingency effects in nonword spelling. *Quarterly Journal of Experimental Psychology, 40A*, 5–40.

Baxter, D. M., & Warrington, E. K. (1987). Transcoding sound to spelling: Single or multiple sound unit correspondence? *Cortex, 23*, 11–28.

Beauvois, M. F., & Dérouesné, F. (1981). Lexical or orthographic agraphia. *Brain, 104*, 21–49.

Bi, Y., Han, Z., Shu, H., & Weekes, B. (2007). The interaction between the semantic and the nonsemantic routes of reading: Evidence from Chinese. *Neuropsychologia, 45*, 2660–2673.

Bird, H., Howard, D., & Franklin, S. (2003). Verbs and nouns: The importance of being imageable. *Journal of Neurolinguistics, 16*, 113–149.

Bonin, P., Barry, C., Méot, A., & Chalard, M. (2004). The influence of age of acquisition in word reading and other tasks: A never ending story? *Journal of Memory and Language, 50*, 456–476.

Bonin, P., & Méot, A. (2002). Writing to dictation in real time in adults: What are the determinants of written latencies? In S. P. Shohov (Ed.), *Advances in psychology research* (Vol. 16). New York: Nova Science Publishers.

Bonin, P., Méot, A., Mermillod, M., Ferrand, L., & Barry, C. (2009). The effects of age of acquisition and frequency trajectory on object naming: Comments on Pérez (2007). *Quarterly Journal of Experimental Psychology, 62*, 1132–1140.

Bosse, M. L., Valdois, S., & Tainturier, M. J. (2003). Analogy without priming in early spelling development. *Reading and Writing, 16*, 693–716.

Brysbaert, M., Van Wijnendaele, I., & De Deyne, S. (2000). Age-of-Acquisition of words is a significant variable in semantic tasks. *Acta Psychologica, 104*, 215–216.

Bub, D., & Kertesz, A. (1982). Deep agraphia. *Brain and Language, 17*, 146–165.

Cambell, R. (1983). Writing nonwords to diction. *Brain and Language, 19*, 153–178.

Cuetos, F., & Barbon, A. (2006). Word naming in Spanish. *European Journal of Cognitive Psychology, 18*, 415–436.

Delattre, M., Bonin, P., & Barry, C. (2006). Written spelling to dictation: Sound-to spelling regularity affects both writing latencies and durations. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 32*, 1330–1340.

Duhamel, J. R., & Poncet, M. (1986). Deep dysphasia in a case of phonemic deafness: Role of the right hemisphere in auditory language comprehension. *Neuropsychologia, 24*, 769–779.

Fiser, J., & Aslin, R. N. (2002). Statistical learning of new visual feature combinations by infants. *Proceedings of the National Academy of Sciences of the Unites States of America, 99*, 15822–15826.

Folk, J. R., & Jones, A. C. (2004). The purpose of lexical/sublexical interaction in spelling: Evidence from dysgraphia. *Neurocase, 10*, 65–69.

Forster, K. I., & Forster, J. C. (2003). Dmdx: A windows display program with millisecond accuracy. *Behavior Research Methods, Instruments, and Computers, 35*, 116–124.

Gao, S., Wang, Y., Shi, S., Liu, J., Lin, G., Rao, B., et al. (1993). *Aphasia*. Beijing: Union Press of Beijing Medical University, Beijing Union Medical College.

Gerhand, S., & Barry, C. (1998). Word frequency effects in oral reading are not merely age-of-acquisition effects in disguise. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 24*, 267–283.

Hillis, A. E., & Caramazza, A. (1991). Mechanisms for accessing lexical representations for output: Evidence from a category-specific semantic deficit. *Brain and Language, 40*, 106–144.

Hillis, A. E., & Caramazza, A. (1995). Converging evidence for the interaction of semantic and sublexical phonological information in accessing lexical representations for spoken output. *Cognitive Neuropsychology, 12*, 187–227.

Hillis, A. E., Rapp, B., & Caramazza, A. (1999). When a rose is a rose in speech but a tulip in writing. *Cortex, 35*, 337–356.

Houghton, G., & Zorzi, M. (2003). Normal and impaired spelling in a connectionist dual-route architecture. *Cognitive Neuropsychology, 20*, 115–162.

Kim, H., & Na, D. I. (2000). Dissociation of pure Korean words and Chinese-derivative words in phonological dysgraphia. *Brain and Language, 74*, 134–137.

Kirkham, N. Z., Slemmer, J. A., & Johnson, S. P. (2002). Visual statistical learning in infancy: evidence for a domain general learning mechanism. *Cognition, 83*, B35–B42.

Laiacona, M., Capitani, E., Zonca, G., Scola, I., Saletta, P., & Luzzatti, C. (2009). Integration of lexical and sublexical processing in the spelling of regular words: A multiple single-case study in Italian dysgraphic patients. *Cortex, 45*, 804–815.

Law, S. (2004). Writing errors of a cantonese dysgraphic patient and their theoretical implications. *Neurocase, 10*, 132–140.

Law, S., Wong, W., & Kong, A. (2006). Direct access from meaning to orthography in Chinese: A case study of superior written to oral naming. *Aphasiology, 20*, 565–578.

Law, S. P., & Or, B. (2001). A case study of acquired dyslexia and dysgraphia in Cantonese: Evidence for nonsemantic pathways for reading and writing in Chinese. *Cognitive Neuropsychology, 18*, 729–748.

Law, S. P., Yeung, O., Wong, W., & Chiu, K. M. Y. (2005). Processing of semantic radicals in writing Chinese characters: Data from a Chinese dysgraphic patient. *Cognitive Neuropsychology, 22*, 885–903.

Lee, C. Y., Huang, H. W., Kuo, W. J., Tsai, J. L., & Tzeng, O. J. (2010). Cognitive and neural basis of the consistency and lexicality effects in reading Chinese. *Journal of Neurolinguistics, 23*, 10–27.

Lee, C. Y., Tsai, J. L., Chiu, Y. C., Hung, D. L., & Tzeng, O. J. (2006). The early extraction of sublexical phonology in reading Chinese pseudocharacters: An event-related potentials study. *Language and Linguistics, 7*, 619–636.
Lee, C. Y., Tsai, J. L., Kuo, W. J., Yeh, T.C., Wu, Y. T., Ho, L. T., et al. (2004). Neuronal correlates of consistency and frequency effects on Chinese character naming: An event-related fMRI study. *NeuroImage, 23*, 1235–1245.
Lété, B., Peereman, R., & Fayol, M. (2008). Consistency and word-frequency effects on spelling among first- to fifth-grade French children: A regression-based study. *Journal of Memory and Language, 58*, 952–977.
Lewis, M. B., Gerhand, S., & Ellis, H. D. (2001). Re-evaluating age-of-acquisition effects: Are they simply cumulative-frequency effects? *Cognition, 78*, 189–205.
Liu, Y., Shu, H., & Li, P. (2007). Word naming and psycholinguistic norms: Chinese. *Behavior Research Methods, 39*, 192–198.
Luzzi, S., Bartolini, M., Coccia, M., Provinciali, L., Piccirilli, M., & Snowden, J. S. (2003). Surface dysgraphia in a regular orthography: Apostrophe use by an Italian writer. *Neurocase, 9*, 285–296.
Majerus, S., Lekeu, F., Van der Linden, M., & Salmon, E. (2001). Deep dysphasia: Further evidence on the relationship between phonological short term memory and language processing. *Cognitive Neuropsychology, 18*, 385–410.
Miceli, G., Benvegnù, B., Capasso, R., & Caramazza, A. (1997). The independence of phonological and orthographic lexical forms: Evidence from aphasia. *Cognitive Neuropsychology, 14*, 35–70.
Miceli, G., & Capasso, R. (2006). Spelling and dysgraphia. *Cognitive Neuropsychology, 23*, 110–134.
Miceli, G., Capasso, R., Ivella, A., & Caramazza, A. (1997). Acquired dysgraphia in alphabetic and stenographic writing. *Cortex, 33*, 355–367.
Nolan, K. A., & Caramazza, A. (1983). An analysis of writing in a case of deep dyslexia. *Brain and Language, 20*, 305–328.
Norton, E. S., Kovelman, I., & Petitto., L. A. (2007). Are there separate neural systems for spelling? New insights into the role of rules and memory in spelling from fMRI. *International Journal of Mind, Brain, and Education, 1*, 1–12.
People's Education Press. (2006). *Standard textbooks for nine-year elementary to high school education*. Beijing: Author.
Pérez, M. A. (2007). Age-of-Acquisition persists as the main factor in picture naming when cumulative word-frequency and frequency trajectory are controlled. *Quarterly Journal of Experimental Psychology, 60*, 32–42.
Rapp, B., Epstein, C., & Tainturier, M. J. (2002). The integration of information across lexical and sublexical processes in spelling. *Cognitive Neuropsychology, 19*, 1–29.
Reich, S., Chou, T. L., & Patterson, K. (2003). Acquired dysgraphia in Chinese: Further evidence on the links between phonology and orthography. *Aphasiology, 17*, 585–604.
Sadoski, M., Willson, V. L., Holcomb, A., & Boulware-Gooden, R. (2004). Verbal and nonverbal predictors of spelling performance. *Journal of Literacy Research, 36*, 461–478.
Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996). Statistical learning by 8-month-old infants. *Science, 274*, 1926–1928.
Saffran, J. R., Pollak, S. D., Seibel, R. L., & Shkolnik, A. (2007). Dog is a dog is a dog: Infant rule learning is not specific to language. *Cognition, 105*, 669–680.
Shallice, T. (1981). Phonological agraphia and the lexical route in writing. *Brain, 104*, 413–429.
Shen, H. H., & Bear, D. R. (2000). Development of orthographic skills in Chinese children. *Reading and Writing, 13*, 197–236.
Shu, H., Chen, X., Anderson, R., Wu, N., & Xuan, Y. (2003). Properties of school Chinese. *Child Development, 74*, 27–47.
Shu, H., & Wu, N. (2006). Growth of orthography–phonology knowledge in the Chinese writing system. In P. Li, L. H. Tan, E. Bates, & O. J. L. Tzeng (Eds.), *Handbook of east Asian psycholinguistics* (pp. 103–113). Cambridge: Cambridge University Press.
Standards Press of China. (1994). *Information technology—UCS: Universal multiple-octet coded character set: Part 1. Architecture and basic multilingual plane*. Beijing: Author.
Steyvers, M., & Tenenbaum, J. B. (2005). The large-scale structure of semantic networks: Statistical analysis and a model of semantic growth. *Cognitive Science, 29*, 41–78.

Sun, H. L., Huang, J. P., Sun, D. J., Li, D. J., & Xing, H. B. (1997). Introduction to language corpus system of modern Chinese study. In M. Y. Hu (Ed.), *Paper collection for the 5th World Chinese Teaching Symposium.* Beijing: Peking University Publisher.

Tainturier, M. J., & Caramazza, A. (1996). The status of double letters in graphemic representations. *Journal of Memory and Language, 35*, 53–73.

Valdois, S., Carbonnel, S., David, D., Rousset, S., & Pellat, J. (1995). Confrontation of PDP models and dual-route models through the analysis of a case of deep dysphasia. *Cognitive Neuropsychology, 12*, 681–724.

Ward, J., Stott, R., & Parkin, A. J. (2000). The role of semantics in reading and spelling: Evidence for the "summation hypothesis." *Neuropsychologia, 38*, 1643–1653.

Weekes, B. S., & Raman, I. (2008). Bilingual deep dysphasia. *Cognitive Neuropsychology, 25*, 411–436.

Weekes, B. S., Yin, W., Su, I. F., & Chen, M. J. (2006). The cognitive neuropsychology of reading and writing in Chinese. *Language and linguistics, 7*, 595–617.

Xing, H., Shu, H., & Li, P. (2004). The acquisition of Chinese characters: Corpus analyses and connectionist simulations. *Journal of Cognitive Science, 5*, 1–49.

Zevin, J. D., & Seidenberg, M. S. (2002). Age of acquisition effects in word reading and other tasks. *Journal of Memory and Language, 47*, 1–29.

Ziegler, J. C., & Ferrand, L. (1998). Orthography shapes the perception of speech: The consistency effect in auditory word recognition. *Psychonomic Bulletin & Review, 5*, 683–689.