



Surface dyslexia in a Japanese patient with semantic dementia: evidence for similarity-based orthography-to-phonology translation

Takao Fushimi^{a,*}, Kenjiro Komori^b, Manabu Ikeda^b, Karalyn Patterson^c,
Mutsuo Ijuin^a, Hiroataka Tanabe^b

^a Language, Cognition, Brain Science Research Group, Tokyo Metropolitan Institute of Gerontology, 35-2 Sakae-cho, Itabashi-ku, Tokyo 173-0015, Japan

^b Department of Neuropsychiatry, Ehime University School of Medicine, Shitsukawa, Shigenobu-cho, Onsen-gun, Ehime 791-0295, Japan

^c MRC Cognition and Brain Sciences Unit, 15 Chaucer Road, Cambridge CB2 2EF, UK

Received 19 November 2001; received in revised form 26 June 2002; accepted 7 April 2003

Abstract

We studied the reading performance of a Japanese-speaking patient, TI, with bilateral but asymmetrical (left more than right) temporal-lobe atrophy, severe anomia, and poor word comprehension. Most Japanese kanji characters correspond to several different legitimate pronunciations in different contexts, with varying degrees of correspondence consistency. TI made many errors in reading aloud words that violate statistically typical character-sound correspondences, especially for less common words, but had relatively preserved ability to read aloud strings in which the assignment of the typical pronunciation for each component character yields the correct pronunciation for the whole word. The degree of consistency of character-sound correspondences affected his performance on both words and nonwords in a graded manner. One interpretation is that TI's surface dyslexic reading reflects intact direct computation of phonology from orthography, but without the additional constraint from word meaning that is, in this framework, considered critical for correct pronunciation of lower-frequency words with atypical character-sound correspondences. Another interpretation is that TI's performance reflects partially damaged lexical knowledge of whole-word orthography and phonology, coupled with spared sublexical knowledge of character-sound correspondence rules. Whichever of these interpretations is preferred, this study offers the most detailed information yet available on the characteristics of surface dyslexia in Japanese.

© 2003 Published by Elsevier Ltd.

Keywords: Reading aloud; Japanese kanji; Consistency; Frequency; The dual-route model; The triangle model

1. Introduction

Acquired surface dyslexia as observed in English is a disorder of reading in which patients are impaired at reading aloud words that violate typical spelling-sound correspondences (e.g. exception words like PINT). In contrast, these patients have relatively, or sometimes even perfectly, preserved ability to read aloud letter strings in which the assignment of the typical pronunciation for each component yields the correct pronunciation for the whole string (e.g. regular words like HINT and nonwords like RINT). The usual incorrect reading response produced by surface dyslexic patients has been characterized as a Legitimate Alternative Reading of Components (LARC) error (Patterson, Suzuki, Wydell, & Sasanuma, 1995), in which the pronunciation of one or more components is inappropriate for the target word but is nonetheless legitimate, and indeed usually more typical, for

that component in other words in the target's orthographic neighborhood (e.g. reading PINT as /PINT/ to rhyme with HINT). Performance on words with atypical spelling-sound correspondences is much more compromised for low- than high-frequency words. Surface dyslexia often (Breedin, Saffran, & Coslett, 1994; Funnell, 1996; Graham, Patterson, & Hodges, 1994, 2000; Patterson & Hodges, 1992; Ward, Stott, & Parkin, 2000), although not invariably (Cipolotti & Warrington, 1995; Gerhand, 2001; Lambon Ralph, Ellis, & Franklin, 1995), accompanies semantic impairment, as typically observed in semantic dementia, a progressive, relatively selective impairment of semantic memory associated with atrophy of the anterior, inferior temporal lobes (Chan et al., 2001; Galton et al., 2001; Hodges, Patterson, Oxbury, & Funnell, 1992; Snowden, Goulding, & Neary, 1989).

1.1. Neighborhood consistency effects

Surface dyslexic reading errors have most often been interpreted as the manifestation of subword-level

* Corresponding author. Tel.: +81-3-3964-3241; fax: +81-3-3579-4776.
E-mail address: fushimi@tmig.or.jp (T. Fushimi).

orthography-to-phonology translation on the basis of a set of pronunciation rules like grapheme–phoneme correspondence (GPC) rules; word-specific knowledge of pronunciations indispensable for exception-word reading is assumed to be impaired in a frequency-modulated fashion (e.g. Coltheart, 1985). A few studies, however, have noted that a simple dichotomy between regular versus exception words is not sufficient to describe surface dyslexic reading. Shallice, Warrington, & McCarthy (1983) chose English exception words that they characterized as either mildly irregular (e.g. CROW) or very irregular (e.g. YACHT), and reported that a surface dyslexic patient achieved best performance for regular words (74%), intermediate for mildly regular words (49%), and worst performance for very irregular words (29%). Patterson and Behrmann (1997) demonstrated the impact of degrees of consistency in spelling-sound correspondences across orthographic neighborhoods. The critical neighborhood for pronunciation of monosyllabic English words is usually, though not exclusively, defined in terms of word body, comprising the orthographic vowel and terminal consonant(s), like $_INT$ in HINT, PINT, and so forth (Treiman, Mullennix, Bijeljac-Babic, & Richmond-Welty, 1995). The authors prepared two sets of quartets consisting of a regular word, an exception word and two nonwords sharing the same body. In one set labeled $REG > EXC$ (e.g. HINT, PINT, RINT, and NINT), regular words have more friendly neighbors with the same pronunciation for the shared body than enemies with a discrepant pronunciation for the shared body. In the other set labeled $REG \leq EXC$ (e.g. COST, POST, FOST, and BOST), regular words have no more friends than enemies. The subject of this study was a surface dyslexic patient MP with severe semantic impairment (Behrmann & Bub, 1992). Between these two conditions, $REG > EXC$ versus $REG \leq EXC$, MP's accuracy of reading aloud decreased for both regular words (from 100 to 83%) and nonwords (from 90 to 78%, where correct was defined as the 'regular' pronunciation for a nonword) but increased for exception words (from 38 to 57%). The dual-route theory of word reading (e.g. Coltheart, 1985) assumes that exception-word reading and nonword reading are accomplished with respect to separate lexical and nonlexical knowledge, respectively, and predicts that the preserved nonlexical procedure in surface dyslexia will result in equally efficient performance for regular words and nonwords whether they are from $REG > EXC$ or $REG \leq EXC$ neighborhoods. The consistency effects, which had a parallel impact on MP's reading performance both for nonwords (as well as regular words) and exception words, therefore call into question the standard assumption of independence between the lexical and nonlexical procedures, and instead suggest common knowledge that applies to pronunciation of all letter strings.

In research on reading processes in alphabetic writing systems like English, where various models of reading aloud have been proposed (Coltheart, Curtis, Atkins, & Haller, 1993; Coltheart, Rastle, Perry, & Ziegeler, 2001; Plaut, 1997; Plaut, McClelland, Seidenberg, & Patterson, 1996;

Rastle & Coltheart, 1999; Seidenberg & McClelland, 1989; Seidenberg, Plaut, Petersen, McClelland, & McRae, 1994), these kinds of fine-grained analyses of surface dyslexic reading are likely to provide substantial constraints on the further development of the models. Empirical studies offering similarly fine-grained analyses in nonalphabetic writing systems like Japanese are relatively scarce (Nakamura, Nakanishi, Hamanaka, Nakaaki, & Yoshida, 2000; Patterson et al., 1995), and there has been only one attempt to model relevant performance of Japanese readers (Ijuin, Fushimi, Patterson, & Tatsumi, 1999). With the goal of advancing an understanding of reading processes in Japanese, the present study investigates graded effects of neighborhood consistency on reading performance of a Japanese-speaking patient with surface dyslexia.

1.2. Japanese phonology and orthography

The spoken form of a Japanese word comprises a sequence of from one to several morae (e.g. /ni-ho-N/, meaning *Japan*). A mora corresponds to one of the five following phonological entities: a single vowel; a consonant–vowel compound; a palatalized consonant–vowel compound; the nasal coda /N/; or a geminate consonant. Two kinds of scripts are used in the written forms of Japanese words: phonographic kana and morphographic kanji.

Kana is sub-divided into hiragana and katakana. Both hiragana and katakana systems comprise 75 characters, and there is always a pair of hiragana and katakana characters representing the same pronunciation. A kana character usually maps to a single mora (e.g. く /ku/, し /si/, い /i/), and character-sound correspondences in kana are perfectly transparent. Any word can be written in either hiragana or katakana; but in terms of most typical usage, hiragana is used for most function words, the inflections of verbs and adjectives, and some content words, whereas katakana is used for loan words from western languages.

Kanji comprises approximately 6000 characters, with about half of these used in daily life (National Language Research Institute, 1976). Some kanji characters appear as a single-character noun (e.g. 空/sora/ meaning *sky*), as a component of multiple-character kanji nouns (e.g. 空気/kuu-ki/ *air*, 空色/sora-iro/ *sky blue*), and as the stems of kanji-kana compound verbs (e.g. 空く/a-ku/ *become vacant*) and adjectives (e.g. 空しく/muna-si-i/ *fruitless*)¹. Other characters appear only in one, two, or three of these written word classes. A kanji character does not reliably map to any specific size of phonological unit like a phoneme, a mora, a syllable, or a word, and cannot be decomposed into elements which correspond to phonemes or morae. The majority of kanji characters have two or more legitimate pronunciations, with the one pronunciation appropriate for the character in any given

¹ Among kanji words, two-character words (about 80%) predominate over those with other numbers of characters (Amano & Kondo, 1999; National Language Research Institute, 1970).

word determined by intraword context. The pronunciations of a kanji character can be divided into ON pronunciation adopted from Chinese and KUN pronunciation derived from original Japanese. A kanji character represents not only the morpheme's phonology but also its semantic domain.

1.3. Surface dyslexia in Japanese-speaking cases

Although such analogies are inevitably an oversimplification, there is a sense in which kana words, kana nonwords, and kanji words, respectively, parallel regular words, nonwords, and exception words in English. According to this simplified formulation, surface dyslexia in Japanese-speaking cases should manifest as a selective, frequency-modulated deficit in reading aloud kanji words, coupled with preserved ability to read aloud kana words and nonwords. The dominant pattern of erroneous responses for kanji words should be LARC errors, where patients produce inappropriate but legitimate pronunciations for constituent characters in a word, for example, erroneous assignment of ON versus KUN. Several cases with cerebrovascular accident have been reported to demonstrate this kind of dyslexic pattern, although LARC errors were infrequent (e.g. [Sasanuma, 1980](#); [Sasanuma, 1985](#)). LARC errors were apparently more frequently but only anecdotally observed in patients described as having 'Gogi' (word meaning) aphasia, with profound anomia and poor word comprehension ([Tanabe et al., 1992](#); [Tanabe et al., 1996](#)). The difference between these two aetiologies in error patterns, if confirmed by further research, may be explicable in terms of lesion extent/location and consequent range of language deficits. Patients with Gogi aphasia have progressive but circumscribed atrophy of the anterior, inferior temporal lobe associated with relatively selective impairment of lexico-semantic knowledge. By contrast, surface dyslexia resulting from cerebrovascular accidents is typically associated with perisylvian lesions, often resulting in additional phonological impairments that may yield reading errors other than LARC (e.g. [Sasanuma, 1980](#); [Sasanuma, 1985](#)).

Beyond the kana/kanji dissociation, [Fushimi, Ijuin, Patterson, & Tatsumi \(1999\)](#) concluded that some statistical clues go a long way towards predicting the pronunciation for each component character in kanji words, based on an analysis of a corpus of approximately 31,000 two-character kanji words from a Japanese dictionary ([Nishio, Iwabuchi, & Mizutani, 1986](#)). Consider a kanji character 神 which has four pronunciations, each of which is appropriate for one or two of the following words: a single-character word 神/kami/ god, two-character words 神経/siN-kei/ nerve, 神社/ziN-zja/ shrine, 神様/kami-sama/ call for god, and 神主/kaN-nusi/ priest. Although the pronunciations for 神 vary across these examples, the statistical analysis revealed that, among 86 two-character kanji words sharing 神 at the first position, 神 is pronounced /siN/ in 71 of 86 words, a very high proportion (0.83). This feature of one dominant (or typical) pronunciation shared across the majority of

words containing the character is true of many kanji characters. Indeed, at the extreme of this phenomenon, 40% of the roughly 3000 kanji characters in the corpus analyzed by [Fushimi et al.](#) have only one consistent pronunciation invariant across words sharing the character at the same position (e.g. 満/maN/ and 開/kai/ in 満開/maN-kai/).

These statistical constraints on character pronunciation in kanji words predispose the orthography-to-phonology computation to the correct whole-word pronunciation of words with consistent or typical character-sound correspondences, but would bias the outcome towards an incorrect pronunciation of words with atypical character-sound correspondences. One hypothesis is that (a) direct orthography-to-phonology translation normally operates with reference to knowledge of pronunciations of the target and also of other words within its orthographic neighborhood; (b) the efficiency of this translation (as measured by response times in normal readers) declines as a function of both decreasing word frequency and decreasing consistency (or typicality) of character-sound correspondences; and (c) surface dyslexic reading reflects, in accuracy, the efficiency of this frequency- and similarity-based orthography-to-phonology translation.

[Patterson et al. \(1995\)](#) confirmed this prediction in a striking case of Japanese surface dyslexia, patient NK with marked left-temporal atrophy, profound anomia and impaired word comprehension. Her performance in reading kana strings was flawless both for words and nonwords. In reading aloud two-character kanji words, NK demonstrated a prominent deficit for less common words with atypical character-sound correspondences, coupled with relatively preserved reading performance for common words and/or those with typical character-sound correspondences². The majority of her reading errors were LARC. [Nakamura et al. \(2000\)](#) also presented two other patients with semantic dementia (YK and KI) showing this dyslexic pattern.

This article presents another Japanese-speaking case, TI, with semantic dementia and surface dyslexia. In a significant advance beyond the two studies described above, the stimulus materials for the current experiments (a) employed statistically based, graded degrees of consistency of character-sound correspondences, and (b) included kanji nonwords as well as words. On the assumption that dichotomous effects of typicality observed in previous studies ([Nakamura et al., 2000](#); [Patterson et al., 1995](#)) are not sufficient to capture the implications of surface dyslexic performance for models of reading in Japanese, it is of

² The definition of typicality in [Patterson et al. \(1995\)](#) was based on the assumption that, in two-character kanji words, an ON pronunciation is typically, though not always, correct for each constituent character. Recently, [Amano and Kondo \(1999\)](#) confirmed this assumption: among about 32,000 two-character kanji words in their corpus, those in which an ON pronunciation is correct for each constituent character (85%) substantially predominate over those in which the KUN pronunciation is correct for each constituent character (11%). In contrast, out of 2200 single-character kanji words in their corpus, KUN (60%) slightly predominates over ON (40%) as the correct pronunciation.

substantial interest to determine whether surface dyslexic reading performance for kanji words and nonwords is modulated by the degree of consistency in character-sound correspondences. TI, a patient at a mild-to-moderate stage of semantic dementia, is an ideal case for the exploration of this issue.

2. Case report

2.1. Background and neuroradiological findings

TI, a right-handed 57-year-old office worker with a university education, presented in August 1999 with a 3-year history of progressive language problems including word-finding difficulty, poor word comprehension, difficulty in writing kanji words, and decrease of spontaneous speech, accompanied by mild personality change. He complained that he could understand so few things as not to speak. When he did speak, however, his anomie spontaneous speech was fluent with normal articulation, phonology, and syntax. No salient deterioration was observed in his day-to-day episodic or autobiographical memory.

Magnetic resonance imaging (MRI) in November 1999 (Fig. 1) revealed circumscribed atrophy in the left anterior temporal lobe, and a SPECT scan in September 1999 showed hypometabolism in the same region. An MRI scan in December 2000 revealed that the atrophy had expanded to include left frontal and right anterior temporal regions.

2.2. General neuropsychological assessments

2.2.1. General test batteries

The following assessments were administered in August 1999 after informed consent was obtained from TI and his wife. TI scored 28/30 in the Mini-Mental State Examination (Folstein, Folstein, & McHugh, 1975), 0.5 on the Clinical Dementia Rating Scale (Hughes, Berg, Danziger, Coben, & Martin, 1982), and 18/70 in the Alzheimer Disease Assessment Scale (Rosen, Mohs, & Davis, 1984). On the Japanese version of the Wechsler Adult Intelligence Scale Revised (Shinagawa, Kobayashi, Fujita, & Maekawa, 1990;

Wechsler, 1981), he showed more marked deterioration in verbal (VIQ = 71) than performance IQ (PIQ = 91) as in other cases with semantic dementia (Hodges et al., 1992). Nonlinguistic reasoning ability was well preserved as indicated in a 35/36 score on Raven's Coloured Progressive Matrices (Raven, 1965). Short-term verbal memory was intact (digit span = 7 forward, 5 backward). Visuospatial skills were spared as indicated in a 36/36 score for copy of the Rey complex figure (Rey, 1941) administered in October 2000.

2.2.2. A standard language test

TI's language function was evaluated by the Japanese Standard Language Test for Aphasia (SLTA) (SLTA Committee, 1997), comprising 26 subtests, each assigned to one of five linguistic functions (listening, speaking, reading, writing, and calculating). Table 1 presents TI's performance on a subset of the SLTA. In September 1999, TI scored below two standard deviations from the mean of nonaphasic controls on picture naming of nouns and verbs, scene description, category fluency, written naming of pictures, and writing to dictation with kanji characters. His score for writing the same words to dictation with kana characters was perfect. TI's comprehension performance in this battery was unimpaired for spoken, kanji, and kana word-to-picture matching. Repetition was flawless both for single words and short sentences. At this stage, therefore, TI had only a mild comprehension disorder but severe anomia (also kanji dysgraphia), a pattern which might be characterized as progressive fluent aphasia. As expected, follow-up testing (14 months later in November 2000) revealed deteriorated performance in a number of subtests, including comprehension of spoken and written sentences and commands, reading aloud kanji words, and written picture naming with kana characters. Nevertheless, TI's performance in single-word comprehension on the SLTA was still perfect at this later stage. As indicated by the more demanding task described next, this result is probably explained by the fact that the word-to-picture matching subtests in this battery are insensitive to mild degrees of semantic impairment, because all materials in these subtests are of extremely high familiarity and because the distractors are not always semantically

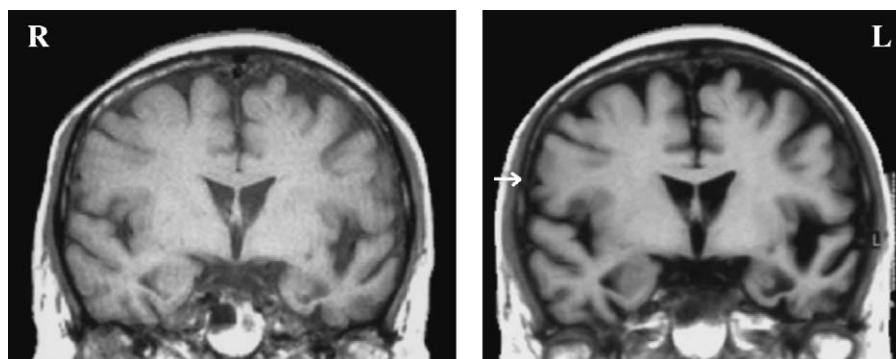


Fig. 1. Coronal sections of an MRI scan for TI in November 1999 (left) and December 2000 (right).

Table 1
 TI's scores on a subset of Japanese Standard Test for Aphasia (SLTA)

| | September 1999 | November 2000 | Control | Full score |
|------------------------------------|----------------|----------------|------------|------------|
| Spoken word-to-picture matching | 10 | 10 | 10.0 ± 0.2 | 10 |
| Kanji word-to-picture matching | 10 | 10 | 9.9 ± 0.8 | 10 |
| Kana word-to-picture matching | 10 | 10 | 10.0 ± 0.1 | 10 |
| Spoken sentence comprehension | 10 | 7 ^a | 9.5 ± 0.8 | 10 |
| Written sentence comprehension | 9 | 6 ^a | 9.4 ± 1.5 | 10 |
| Spoken command comprehension | 9 | 3 ^a | 9.6 ± 0.7 | 10 |
| Written command comprehension | 9 | 6 ^a | 9.4 ± 1.5 | 10 |
| Picture naming (noun) | 8 ^a | 6 ^a | 19.6 ± 0.8 | 20 |
| Picture naming (verb) | 8 ^a | 7 ^a | 9.9 ± 0.4 | 10 |
| Scene description | 4 ^a | 3 ^a | 5.8 ± 0.6 | 6 |
| Word fluency (animal within 1 min) | 3 ^a | 0 ^a | 12.6 ± 4.5 | – |
| Word repetition | 10 | 10 | 10.0 ± 0.1 | 10 |
| Sentence repetition | 5 | 4 | 4.5 ± 0.8 | 5 |
| Reading aloud kanji words | 5 | 4 ^a | 5.0 ± 0.4 | 5 |
| Reading aloud kana words | 5 | 5 | 5.0 ± 0.1 | 5 |
| Written naming with kanji | 1 ^a | 3 | 4.2 ± 1.1 | 5 |
| Written naming with kana | 4 | 3 ^a | 4.8 ± 0.7 | 5 |
| Writing to dictation with kanji | 2 ^a | 5 | 4.3 ± 1.0 | 5 |
| Writing to dictation with kana | 5 | 5 | 4.8 ± 0.8 | 5 |

^a The TI's performance was below two standard deviations from the mean of nonaphasic controls.

related to the target. TI's deficit in written picture naming and writing to dictation with kanji characters improved between 1999 and 2000 because a subset of the test materials had been included in daily therapy of picture naming and written naming with kanji characters.

2.2.3. Comprehension and naming tests of single words

TI's performance in both comprehension and naming of single items was further investigated with materials comprising 80 line drawings and 10 colours, prepared by Ito et al. (1994) for patients characterized as word meaning ('Gogi' in Japanese) aphasics. These items represent 10 exemplars from each of four categories of natural kinds (body parts, colours, animals, and vegetables/fruits) and five categories of artifact (vehicles, sports, utensils, musical instruments, and processed foods). In a spoken word-to-picture matching test, the participant is given a spoken word and asked to point to the target picture presented in an array containing the target plus nine within-category distracters. In a naming test, the participant is asked to name each of 90 pictures (80 line drawings, 10 colours) presented individually. Ito et al. (1994) reported that the performance of normal controls was near ceiling for both tasks and all categories (mean > 98%; S.D. = 2–6%) except naming of musical instruments (mean = 91%; S.D. = 11%). Patients with Gogi aphasia showed worse performance than normal controls in both comprehension and naming tests for all categories except body parts and colours, for which some patients performed within the normal range.

Because word frequency and/or familiarity was not equated either across the nine categories or between natural kinds and artifacts, for the purposes of analysis we re-arranged all materials on the basis of word frequency. Values for printed word frequency (in log₁₀-transformed

occurrences per one million tokens) and spoken word familiarity (from a rating score on a 7-point scale) for each word were obtained from the NTT psycholinguistic database (Amano & Kondo, 1999, 2000). Category exemplars were split into higher- and lower-frequency bands at the median values. Log₁₀-transformed mean (±S.D.) word frequency was 2.63 (±0.40) and 1.60 (±0.53), respectively, for the higher- and lower-frequency bands, while mean (±S.D.) familiarity was 6.20 (±0.32) and 6.07 (±0.32).

TI's performance in September 1999 was clearly below the normal range for comprehension (81%) as well as naming (37%). Out of 57 naming errors, 49 were "do not know" responses, six were circumlocutions, and two were the names of semantically related items. He also demonstrated a remarkable task-by-frequency interaction, with accuracy of 91% versus 71% in the comprehension test for the higher- and lower-frequency bands, respectively, and 64% versus 9% in naming.

2.3. Summary

TI's deficit in single-word comprehension was prominent only for low-frequency words, as in other cases of relatively early semantic dementia or progressive fluent aphasia (Graham, Patterson, & Hodges, 1995). On the other hand, again as in other cases, TI demonstrated profound anomia, with an even more dramatic modulation by word frequency. This discrepancy between degrees of impairment in naming versus comprehension (as measured by word-to-picture matching) is sometimes interpreted as indicating an additional post-semantic deficit in activating phonology. Lambon Ralph, McClelland, Patterson, Galton, and Hodges (2001), however, recently reported that, in semantic dementia, the discrepancy between naming and word-to-picture matching

is greater for patients with left- than right-dominant atrophy, especially at the early and middle stages of progression. Longitudinal assessment reveals that, even for left-dominant patients, whose syndrome at the early stage might be described as progressive pure anomia rather than semantic dementia (Graham et al., 1995; Papagno & Capitani, 2001), a profound comprehension deficit is revealed at later stages, resulting in a typical pattern of semantic dementia. Given that semantic deterioration will invariably harm naming more than word-to-picture matching because of the substantial difference in task demands, we adopt the interpretation of Lambon Ralph et al. (2001) and (a) attribute TI's anomia to an underlying and increasing semantic degradation, and (b) predict that measures of his comprehension will demonstrate much more dramatic impairment at a subsequent stage. We acknowledge, however, that one cannot rule out the possibility that TI, and other left-dominant patients with semantic dementia at a similar stage, have dual deficits in both the semantic system itself and its communication with phonological representations for speech.

To summarize the language performance of TI, he demonstrated severe anomia and moderately impaired single-word comprehension, coupled with preserved articulation and phonology, thus fulfilling the previously reported linguistic criteria for semantic dementia (Hodges et al., 1992; Snowden et al., 1989).

3. Reading performance

3.1. Experiment 1: reading aloud kanji words

For an exploration of predicted neighborhood consistency effects on TI's reading, we adopted the statistically-based definition of consistency of character-sound correspondences suggested by Fushimi et al. (1999). For a certain word (e.g. 神経/siN-kei/), orthographic neighbors were defined as all words, in the corpus of approximately 31,000 two-kanji character words from a Japanese dictionary (Nishio et al., 1986), sharing the same character at the first or the second position (e.g. 神社 and 神様 at the first character; 政経 and 写経 at the second character). As mentioned before, the most typical pronunciation for each constituent character was determined by counting the number of neighbors sharing the same pronunciation for the shared character (e.g. /siN/ for 神 and /kei/ for 経). On the basis of these statistics, two-character kanji words were assigned to three categories: (i) *consistent* if each constituent character has the identical pronunciation across neighbors; (ii) *inconsistent-typical* if each constituent has more than one legitimate pronunciation across neighbors but it is the statistically most typical pronunciation of each character that is appropriate to this target word; (iii) *inconsistent-atypical* if each constituent has more than one legitimate pronunciation and one or both character pronunciations appropriate to this target word are not the statistically most typical ones.

Another way of describing this manipulation is as follows: for consistent words, all neighbors were friends; for inconsistent-typical words, friends dominated enemies; for inconsistent-atypical words, enemies dominated friends with regard to the first, the second, or both character positions. Thus, consistent, inconsistent-typical and inconsistent-atypical words were located around different points on a single continuum of balance between friends and enemies. These points can be measured by the ratio of friends to neighbors, averaged across the two character positions (hereinafter, *friends/neighbors*).³

We assume that, as mentioned before, (a) the orthography-to-phonology translation in two-character kanji words operates with reference to knowledge of pronunciations of both the target and other words within the target's neighborhood; (b) the efficiency of this translation declines as a function of both decreasing word frequency and decreasing consistency of character-sound correspondence; and (c) a surface dyslexic pattern of reading reflects the efficiency of this translation. On the basis of these three assumptions, we predicted that TI's reading performance should be most accurate for consistent words, intermediate for inconsistent-typical words, and most error-prone for inconsistent-atypical words, with the consistency effect more remarkable for low-frequency words.

3.1.1. Methods

Experimental stimuli comprised 120 two-character kanji words, with 20 words in each of the six conditions formed by crossing three bands of consistency with two bands of word frequency (see Fushimi et al., 1999 for full description of the materials). The upper rows of Table 2 present, for each of the six conditions, mean word length (in number of morae), printed word frequency (in log₁₀-transformed occurrences per approximately one million tokens) (National Language Research Institute, 1970), written word familiarity (mean rating on a 7-point scale) (Amano & Kondo, 1999), character frequency (in log₁₀-transformed occurrences per approximately one million word tokens, averaged across the first and the second characters) (National Language Research Institute, 1970), and friends/neighbors. Mean word length was equated across the six word conditions. Within each frequency band, word frequency and character frequency were equated across the three consistency bands. In each of the three consistency bands, high- and low-frequency words were equated in terms of friends/neighbors. The middle rows

³ In Fushimi et al. (1999), the counting of neighbors was based on only two-character kanji words because the great majority of kanji words consist of two characters and because, when the calculation was done, a reliable large corpus for single-character kanji words was not available. However, the statistics can include all kanji words with any number of characters using the more recent corpus (Amano & Kondo, 1999). Furthermore, we do not claim that this friends/neighbors measure is necessarily the best metric to evaluate consistency. See also related articles for another way to measure consistency used for monosyllabic words in English (Jared, McRae, & Seidenberg, 1990; Rastle & Coltheart, 1999; Treiman et al., 1995).

Table 2

Experiment 1: characteristics of stimulus materials (mean \pm S.D.), examples of words with the correct and alternative pronunciations accompanied by normal readers' accuracy, and proportions of words for which TI produced various types of response, as a function of stimulus condition

| | High-frequency | | | Low-frequency | | |
|------------------------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|
| | Consistent | Inconsistent | | Consistent | Inconsistent | |
| | | Typical | Atypical | | Typical | Atypical |
| Word length | 3.55 \pm 0.51 | 3.65 \pm 0.49 | 3.45 \pm 0.51 | 3.55 \pm 0.51 | 3.40 \pm 0.50 | 3.50 \pm 0.51 |
| Word frequency | 1.77 \pm 0.23 | 1.84 \pm 0.23 | 1.72 \pm 0.22 | 0.76 \pm 0.06 | 0.77 \pm 0.06 | 0.78 \pm 0.06 |
| Word familiarity | 6.12 \pm 0.31 | 6.01 \pm 0.36 | 6.02 \pm 0.47 | 5.38 \pm 0.57 | 5.48 \pm 0.49 | 5.44 \pm 0.60 |
| Character frequency | 2.71 \pm 0.27 | 2.77 \pm 0.23 | 2.78 \pm 0.31 | 2.61 \pm 0.31 | 2.52 \pm 0.32 | 2.46 \pm 0.45 |
| Friends/neighbors | 1.00 \pm 0.00 | 0.74 \pm 0.07 | 0.42 \pm 0.13 | 1.00 \pm 0.00 | 0.71 \pm 0.07 | 0.33 \pm 0.15 |
| Example | 労働 | 楽団 | 場合 | 満開 | 郷土 | 寿命 |
| Correct pronunciation | rou-dou | gaku-daN | ba-ai | maN-kai | kjou-do | zju-mjou |
| Alternative pronunciation | – | raku-doN | zjou-gou | – | gou-tsutsi | su-mei |
| Proportion correct in normal | 1.00 \pm 0.00 | 1.00 \pm 0.01 | 0.99 \pm 0.03 | 1.00 \pm 0.00 | 0.99 \pm 0.02 | 0.96 \pm 0.04 |
| Correct response | 0.95 | 0.85 | 0.70 | 0.85 | 0.70 | 0.40 |
| LARC error | 0.00 | 0.05 | 0.25 | 0.00 | 0.05 | 0.45 |
| Visual error | 0.00 | 0.10 | 0.05 | 0.10 | 0.05 | 0.20 |
| Incomplete response | 0.00 | 0.00 | 0.10 | 0.05 | 0.00 | 0.00 |
| “Do not know” response | 0.05 | 0.00 | 0.05 | 0.00 | 0.20 | 0.20 |

of Table 2 present an example word in each stimulus condition with its correct pronunciation and one incorrect but legitimate alternative pronunciation. The performance of normal young adults in reading aloud these words under time pressure was no less than 99% correct for all stimulus conditions except for low-frequency inconsistent-atypical words (96%) (Fushimi et al., 1999).

3.1.2. Results

TI was tested in January 2001 without any time limit for pronunciation of each word. The lower section of Table 2 presents the proportions of words in each of the six conditions for which TI produced correct responses and various types of error. LARC errors, as explained earlier, are those in which the pronunciation for one or both characters is incorrect for the target word but nonetheless legitimate for that character in other words (e.g. /raku-daN/ for 楽団/gaku-daN/; /zjou-gou/ for 場合/ba-ai/). Visual errors are those in which TI's response corresponded to a kanji string visually similar to the target. Incomplete responses were those in which TI either produced only part of the correct whole-word pronunciation or inserted other material (such as the Japanese word 'and') between the two components of the word. Surface dyslexic patients sometimes produce both the correct and an erroneous response to a single target, successively and in either order (Funnell, 1996; Patterson & Behrmann, 1997); in this analysis, for such occurrences both the correct response and the appropriate error type were tallied. Likewise, if TI produced an erroneous response which could be classified into two or more error types, each of these error types was tallied. Thus, proportions of the five response types displayed in Table 2 occasionally add up to more than 1.0.

TI's overall accuracy was 74%, with performance characterized by a graded consistency effect from consistent to

inconsistent-typical to inconsistent-atypical words, and with this gradient more pronounced for low- than high-frequency words. A similar, but of course inverse, pattern was observed in LARC errors whereas other error types did not show any systematic effect of consistency.

A simultaneous multiple logistic regression analysis on correct responses, with printed word frequency (as a categorical variable), consistency (as a categorical variable), written word familiarity, character frequency, and number of morae as predictors, revealed significant effects of consistency (Wald = 13.135, $P < 0.001$), familiarity (Wald = 10.985, $P < 0.001$), and character frequency (Wald = 17.078, $P < 0.001$). Although there was no reliable difference between consistent and inconsistent-typical words (Wald = 1.683, $P = 0.195$), the difference between inconsistent-typical and inconsistent-atypical words was significant (Wald = 5.037, $P < 0.05$). Because it is difficult to reveal statistical significance between the adjacent consistency bands, consistency effects were further evaluated by another simultaneous multiple regression analysis with the same five variables as predictors but word frequency (in \log_{10} -transformed value) and consistency (in friends/neighbors) as continuous variables. This regression yielded a significant consistency effect (Wald = 7.920, $P < 0.005$), with additional significant effects of familiarity (Wald = 8.546, $P < 0.005$) and character frequency (Wald = 7.860, $P < 0.005$), indicating that the balance of friends and enemies was a reliable variable to predict success or failure of TI's performance in reading aloud two-character kanji words.

3.1.3. Discussion

As predicted, TI demonstrated a graded consistency effect, with performance best for consistent, intermediate for inconsistent-typical, and worst for inconsistent-atypical words, especially for lower-frequency words. These results

are consonant with the hypotheses that TI's performance reflects the efficiency of frequency- and similarity-based orthography-to-phonology translation, which operates with reference to knowledge of pronunciations of the target word and of other words within the target's orthographic neighborhood.⁴ Thus, the simple dichotomous effect of typicality of character-sound correspondences observed in previous studies (Nakamura et al., 2000; Patterson et al., 1995) can be re-considered as just one aspect of a graded effect of statistically defined consistency.

The nature of TI's impairment may be illustrated in the words for which he produced both the correct response and a LARC error. In these five cases, all in the inconsistent-atypical conditions, TI produced initially a LARC error and then a correct pronunciation in three words, and vice versa in two words. Because inconsistent-atypical words in the present stimulus set were not extremely atypical, as indicated in the value of friends/neighbors around 0.4 rather than close to zero, it is reasonable to expect activation of two candidate pronunciations. Our interpretation of these examples is that the direct orthography-to-phonology procedure activates both the correct and an alternative pronunciation, and that this 'choice' for normal readers is resolved by additional word-specific activation of the correct pronunciation by information that is unavailable or insufficient in patients with surface dyslexia. Given that all convincing cases of Japanese-speaking surface dyslexia published to date have demonstrated severe anomia plus semantic impairment, we suggest that this additional boost to activation of the correct pronunciation of the target word in normal readers might come from word meaning.⁵

In Experiment 2, we consider the further implications of TI's performance in reading kanji nonwords for interpretation of his deficit.

3.2. Experiment 2: reading aloud kanji nonwords

Surface dyslexic patients generally demonstrate preserved nonword reading, as observed in both English-speaking cases (Funnell, 1996; McCarthy & Warrington, 1986; Patterson & Behrmann, 1997) and Japanese cases asked to

⁴ It should be emphasised that the orthography-to-phonology translation is not a kind of compensatory process developed by patients with surface dyslexia after the onset of their disease, but the process by which normal reading aloud is achieved. Response times of normal skilled readers in a speeded reading condition, where readers were asked to pronounce an individually presented word as quickly and accurately as possible, demonstrated a significantly graded consistency effect for low-frequency words with the same materials as used for TI in the present study (Fushimi et al., 1999).

⁵ This suggestion does not mean that anomia must always be accompanied by surface dyslexia. Because phonology in reading aloud can be jointly activated by both orthography and semantics, whereas phonology in picture naming must derive only from semantics, mildly reduced activation from semantics to phonology will impair picture naming more than it disrupts reading aloud, even for low-frequency words with atypical print-sound correspondences (Lambon Ralph, Cicolotti, & Patterson, 1999).

read aloud *kana* nonwords (Nakamura et al., 2000; Patterson et al., 1995; Sasanuma, 1980; Sasanuma, 1985). Reading performance for *kanji* nonwords, however, has never been assessed in Japanese patients, perhaps for theoretical and/or methodological reasons. First, the very concept of nonwords comprising morphographic kanji characters might be ambiguous because the component characters themselves are not meaningless. Second, the criteria for correct pronunciation of kanji nonwords might be ambiguous because the majority of kanji characters have multiple legitimate pronunciations. We claim, however, that two-character kanji nonwords can be defined as nonexistent combinations of real kanji characters, and that, given the evidence for similarity-based orthography-to-phonology translation as demonstrated in TI's performance of reading aloud two-character kanji words, the same procedure should apply to these nonwords.

Consider a two-character kanji nonword (e.g. 満送), each constituent character of which has a single invariant pronunciation (viz. /maN/ for 満 and /sou/ for 送) across real word neighbors sharing the same characters at the same position. The pronunciation for this kind of nonword should be a simple concatenation of the pronunciations for the first and the second characters (viz. /maN-sou/). For a nonword (e.g. 神間) comprising characters with multiple pronunciations (e.g. /siN, ziN, kami, kaN/ for 神 and /kaN, keN, ma, aida/ for 間), any concatenation of these legitimate pronunciations for the first and the second characters (e.g. /siN-kaN, ziN-ma/) might be treated as a correct pronunciation. In this nonword 神間, the most typical pronunciation for the first character (/siN/ for 神) is shared by 71 of its 86 real word neighbors (words with 神 at the first position) whereas that for the second character (/ma/ for 間) is shared by only 30 of its 72 real word neighbors (words with 間 at the second position). Furthermore, the next most common pronunciation in the word neighborhood with regard to the second character, /kaN/, applies to 28 of the 72 neighbors. In other words, the distribution of pronunciations across real word neighbors of this nonword is highly biased to the most typical pronunciation for the first character, whereas the most plausible pronunciation for the second character is rather ambiguous. Thus, the degree of typicality of the statistically most typical pronunciation for each constituent character of a two-character kanji nonword can be measured by the number of words with the most typical pronunciation for the character in the real word neighborhood, divided by the number of neighbors. This value (hereinafter typical/neighbors), when averaged across the first and the second character, is a measure of the consistency for a two-character kanji nonword, parallel to friends/neighbors as a measure of consistency for two-character kanji words.

Fushimi et al. (1999) proposed that consistency should be the main factor governing the choice of pronunciation for each constituent character of a two-character kanji nonword. On the basis of the typical/neighbors measure, two-character kanji nonwords were classified into three

categories: (i) *consistent* if each constituent character has a single invariant pronunciation across real word neighbors (i.e. typical/neighbors for this nonword = 1); (ii) *inconsistent-biased* if each constituent character has multiple pronunciations across real word neighbors but the typical/neighbors ratio for this nonword is relatively high; (iii) *inconsistent-ambiguous* if each constituent character has multiple pronunciations across real word neighbors and typical/neighbors for this nonword is relatively low. Thus, consistent, inconsistent-biased and inconsistent-ambiguous nonwords can be considered as located around different points of a single continuum of balance between more and less common pronunciations within real word neighborhoods, measured by typical/neighbors.

Fushimi et al. (1999) reported that, for normal readers asked to read aloud kanji nonwords, assignment of the statistically most typical pronunciation was most frequent for consistent, intermediate for inconsistent-biased, and least frequent for inconsistent-ambiguous nonwords. On the assumption that patients with surface dyslexia read nonwords by the same similarity-based orthography-to-phonology translation used by normal readers, we predicted that TI's nonword performance should be similarly sensitive this manipulation of consistency.

3.2.1. Methods

Experimental stimuli comprised 120 two-character kanji nonwords, with 20 nonwords in each of the six conditions formed by crossing three bands of consistency (consistent, inconsistent-biased, and inconsistent-ambiguous) with two bands of character frequency (high-character-frequency and low-character-frequency). Consistent nonwords were generated from the consistent words used in Section 3.1, by randomly recombining a character at position 1 of a

consistent word with a character at position 2 of another consistent word to constitute a nonreal combination of real characters. Inconsistent nonwords were created in the same way but from inconsistent words (both inconsistent typical and atypical), and then split at the median value of typical/neighbors into two bands, namely, inconsistent-biased nonwords with higher values on the typical/neighbors measure and inconsistent-ambiguous nonwords with lower values on this measure. In each of the three consistency bands created in this way, nonwords were split at the median value of character frequency, averaged across the first and the second character, into high-character-frequency and low-character-frequency bands. (See Fushimi et al., 1999 for further details on these nonword stimuli).

The upper rows of Table 3 present mean character frequency (in log₁₀-transformed occurrences per approximately one million word tokens, averaged across the first and the second characters) (National Language Research Institute, 1970), and the typical/neighbors measure, as a function of stimulus condition. As a result of how these nonwords were created, within each character-frequency band, character frequency was roughly equated across the three consistency bands. Within each consistency band, the typical/neighbors measure was roughly equated across the two character-frequency bands. If any concatenation of legitimate pronunciations of a two-character kanji nonword is treated as a correct response, normal young adults achieved 89% correct performance in reading aloud these nonwords (range 83–93% across the six conditions), even under time pressure (see proportion correct in the middle section of Table 3; values from Fushimi et al., 1999). Of greater relevance here, in all correct responses produced across the 20 normal readers who participated in the experiment, the percentage of responses where participants

Table 3

Experiment 2: Characteristics of stimulus materials (mean ± S.D.), examples of nonwords with the typical and alternative pronunciations accompanied by normal readers' performance, and proportions of nonwords for which TI produced various types of response, as a function of stimulus condition

| | High character frequency | | | Low character frequency | | |
|------------------------------|--------------------------|------------------------|-------------|-------------------------|------------------------|-------------|
| | Consistent | Inconsistent Biased | Ambiguous | Consistent | Inconsistent Biased | Ambiguous |
| Character frequency | 2.89 ± 0.17 | 2.93 ± 0.19 | 2.98 ± 0.14 | 2.43 ± 0.23 | 2.35 ± 0.28 | 2.37 ± 0.26 |
| Typical/neighbors | 1.00 ± 0.00 | 0.76 ± 0.05 | 0.60 ± 0.07 | 1.00 ± 0.00 | 0.77 ± 0.06 | 0.58 ± 0.07 |
| Example | 集学 | 松合 | 神間 | 満送 | 役岸 | 毛西 |
| Typical pronunciation | juu-gaku | matsu-ai | siN-ma | maN-sou | yaku-gaN | ke-sai |
| Alternative pronunciation | – | sjou-gou | kami-kaN | – | eki-kisi | mou-sei |
| Proportion correct in normal | 0.92 ± 0.07 | 0.93 ± 0.07 | 0.88 ± 0.09 | 0.92 ± 0.07 | 0.83 ± 0.11 | 0.85 ± 0.09 |
| Typical/correct in normal | 0.99 ± 0.03 | 0.81 ± 0.07 | 0.41 ± 0.10 | 1.00 ± 0.01 | 0.78 ± 0.12 | 0.39 ± 0.08 |
| Correct response | 0.90 | 0.90 | 0.90 | 0.50 | 0.50 | 0.50 |
| Typical response | 0.90 | 0.70 | 0.30 | 0.50 | 0.35 | 0.20 |
| LARC response | 0.00 | 0.20 | 0.60 | 0.00 | 0.35 | 0.55 |
| Visual error | 0.15 | 0.10 | 0.10 | 0.45 | 0.25 | 0.25 |
| Incomplete response | 0.00 | 0.00 | 0.00 | 0.05 | 0.15 | 0.20 |
| “Do not know” response | 0.00 | 0.00 | 0.00 | 0.00 | 0.05 | 0.00 |
| Other | 0.00 | 0.00 | 0.00 | 0.00 | 0.05 | 0.00 |
| Typical/correct | 1.00 | 0.78 | 0.33 | 1.00 | 0.70 | 0.40 |

assigned the statistically most typical pronunciations for both constituent characters of a target nonword was 99, 79, and 40%, respectively, for consistent, inconsistent-biased, and inconsistent-ambiguous nonwords, collapsed across the two-character frequency bands. The middle section of Table 3 also presents these values (abbreviated as typical/correct) as a function of stimulus condition, with examples of nonwords accompanied by the typical and the alternative pronunciations.

3.2.2. Results

TI was tested in February and March 2001. The lower rows of Table 3 present the proportions of his various response types for each of the six stimulus conditions. Typical responses are those where TI assigned the statistically most typical pronunciations to both constituent characters (e.g. /yaku-gaN/ for 役岸). LARC responses were those where TI's pronunciation was not the statistically most typical pronunciation for one or both characters but was legitimate for the character(s) in real word neighbors. Visual errors, incomplete and "don't know" responses were defined as described in Section 3.1.2. Other responses were those which did not belong to any type of response mentioned above. When TI produced typical and/or LARC responses to both characters, it was counted as a correct whole-nonword pronunciation. When he produced multiple responses or a response belonging to two or more response types, each response type was tallied, resulting in the proportions occasionally adding up to more than 1.0.

TI's accuracy was 70% overall, and was within the normal range for the high-character-frequency band (90%), but poor for the low-character-frequency band (50%), without any consistency effect in either band. In the occurrences of typical responses, however, a graded consistency effect as well as a character-frequency effect emerged. Typical responses occurred most frequently for consistent, at an intermediate level for inconsistent-biased, and least frequently for inconsistent-ambiguous nonwords, with their occurrence more dominant in the high- than the low-character-frequency band. Inversely, LARC responses occurred least frequently for consistent, at an intermediate rate for inconsistent-biased, and most frequently for inconsistent-ambiguous nonwords. Except for incomplete responses in the low-character-frequency band, there was no systematic effect of consistency on other response types, but these erroneous responses were more frequent for low- than high-character-frequency nonwords, resulting in TI's lower accuracy for low-character-frequency nonwords.

The bottom row of Table 3 presents the proportion of typical to correct responses in each of the six stimulus conditions. Although TI made more errors for low-character-frequency nonwords than normal readers, for the nonwords that TI did read correctly, the proportion of responses with typical pronunciation was within the normal range. This pattern demonstrates the predicted pattern of a graded consistency effect on assignment of the statisti-

cally most typical pronunciations; and the typical/correct measure of TI's performance is nearly identical to that of normal readers in each stimulus condition.

Because this ratio score of typical/correct cannot be submitted to a regression analysis in a single case design, the occurrences of typical responses were submitted to a simultaneous logistic regression analysis with character frequency and consistency as categorical predictors, both of which yielded significant effects (consistency: Wald = 16.122, $p < 0.001$; character frequency: Wald = 10.543, $P < 0.001$). Although the difference between consistent and inconsistent-biased nonwords was marginal (Wald = 2.814, $P = 0.093$), the difference between inconsistent-biased and inconsistent-ambiguous nonwords was significant (Wald = 6.714, $P < 0.01$). When the occurrences of typical responses were submitted to another simultaneous logistic regression analysis with character frequency (\log_{10} -transformed values) and consistency (typicals/neighbors ratio) as continuous predictors, both yielded significant contributions (consistency: Wald = 18.666, $P < 0.001$; character frequency: Wald = 15.683, $P < 0.001$).

3.2.3. Discussion

Our predictions were partially confirmed. TI pronounced two-character kanji nonwords composed of high-frequency characters with accuracy similar to normal readers. Both in high- and low-character-frequency bands, he demonstrated a consistency effect on the production of typical pronunciations, graded from consistent to inconsistent-biased to inconsistent-ambiguous nonwords, with proportions of typical to correct responses strikingly similar to those of normal readers.

TI's lower-than-normal accuracy in reading nonwords composed of low-frequency characters disconfirms our prediction that his nonword reading would be normal. Our purpose in testing his nonword reading was, however, more to demonstrate a normal impact of consistency than to establish completely normal levels of accuracy, and this goal is satisfied by the virtual identity of the typical/correct scores for TI and normal readers. The fact that TI made a number of visual errors and incomplete responses to nonwords composed of lower-frequency characters—errors that rarely occur in normal participants' nonword reading and indeed were not so common in TI's own real-word reading—is, however, to be noted; and it would be of interest to establish in future studies with TI (or other Japanese surface dyslexic cases) whether this phenomenon has a primarily orthographic or phonological basis.

3.3. Reading aloud kana words and nonwords

We briefly report here TI's performance in reading aloud kana strings. Kana word stimuli were hiragana transcriptions of 20 two-character kanji words, a subset from those employed in Section 3.1, which should be more precisely labeled pseudohomophones (viz. orthographic nonwords with

a familiar phonological pattern) because these words are normally written in kanji. Kana nonwords were hiragana transcriptions of the typical pronunciation for 20 two-character kanji nonwords, a subset from those used in Section 3.2, which were not homophonic to any real words. TI's reading accuracy was perfect both for kana pseudohomophones and nonwords.

4. General discussion

The general cognitive profile of TI, a Japanese man with focal and progressive atrophy of the anterior temporal lobe (bilateral but left-dominant), corresponds to the consensus criteria for semantic dementia (Hodges et al., 1992; Snowden et al., 1989). In reading, TI demonstrated a striking pattern of surface dyslexia. His performance in reading aloud was perfect for kana words (more precisely kana pseudohomophones) and kana nonwords, which have completely transparent character-sound correspondences. For two-character kanji words with varying degrees of consistency of character-sound correspondences, his success deteriorated as a function of both decreasing word frequency and decreasing consistency of character-sound correspondences, measured by the ratio of friends/neighbors. The majority of his erroneous responses were LARC errors, in which TI assigned the most typical pronunciation to one or both component characters where the correct pronunciation for that character in the target word is not the statistically typical one. TI's ability to read aloud two-character kanji nonwords was normal, at least for nonwords comprising higher-frequency characters. The probability with which he gave statistically typical pronunciations to the component characters of nonwords exactly mirrored that of normal readers, decreasing with advancing ambiguity as measured by the ratio of typical/neighbors.

These results are consistent with our hypotheses that: (a) for normal readers of Japanese morphographic kanji as well as alphabetic English, direct translation from orthography to phonology operates with reference to knowledge of pronunciations of the target and also of other words within its orthographic neighborhood; (b) again for normal readers of both morphographic and alphabetic scripts, the efficiency of this translation declines as a function of both decreasing word frequency and decreasing consistency of character-sound correspondences; and (c) surface dyslexic reading in both kanji and English reflects, in accuracy, the efficiency of this frequency- and similarity-based orthography-to-phonology translation. Evidence for point (b) comes not only from response-time data of normal readers but also the fact that they make occasional LARC errors when required to read aloud low-frequency words with atypical character-sound correspondences under time pressure (Fushimi et al., 1999).

We acknowledge here that the difference between consistent and inconsistent-typical words did not reach conven-

tional levels of statistical reliability in Experiment 1, and that the difference between consistent and inconsistent-biased nonwords was only marginally significant in Experiment 2. We confidently predict, however, that when additional statistical power is available from data for a case series of Japanese patients with surface dyslexia, a highly significant graded effect of consistency will emerge. In the meantime, we hope that the current study will provide a theoretical framework inspiring subsequent studies on reading aloud of Japanese. The question to be addressed now and in the future is this: in a surface dyslexic reader like TI, with essentially normal ability to read aloud kana words, kana nonwords, kanji nonwords composed of high-frequency characters, and higher-frequency kanji real words composed of consistent characters, what is the nature of the deficit that makes him so much more prone than normal readers are to LARC errors when pronouncing inconsistent kanji words? We discuss two alternative interpretations.

4.1. Interpretations by and implications for models of reading aloud

As a starting point, we propose that any plausible model of reading aloud in Japanese must now account for the clear-cut consistency effects on response times in kanji word-reading (and pronunciation assignment probabilities in kanji nonword reading) for normal participants (Fushimi et al., 1999) and likewise the clear-cut consistency effects in kanji reading accuracy demonstrated for TI and other patients with semantic impairment (Nakamura et al., 2000; Patterson et al., 1995). These demonstrations require a major component of kanji processing in which the pronunciation activated for each constituent character of a kanji word or nonword is a graded function of consistency of character-sound correspondences. We further suggest that models developed for reading aloud in alphabetic scripts, especially dual-route and triangle models, may be pertinent to an understanding of consistency effects in Japanese kanji.

Standard dual-route models assume two strictly separate and independent processes for reading aloud: the lexical and the nonlexical routes (e.g. Coltheart, 1985). In the lexical route, the target word activates its stored whole-word orthographic representation which in turn activates the corresponding whole-word phonological representation. This process is accomplished by the lexical-nonsemantic route, with perhaps an additional contribution from the lexical-semantic route, and thus provides correct pronunciations both for regular and exception words, with greater efficiency for high- than low-frequency words. The nonlexical route translates each subword-level orthographic component into its phonological counterpart on the basis of a set of pronunciation rules (e.g. GPC rules) and then concatenates these phonological elements. It thus provides a rule-governed pronunciation which is correct for regular words and also nonwords but erroneous for exception words. In this framework, substantial but incomplete damage to the frequency-modulated

lexical route, coupled with a spared nonlexical route, produces surface dyslexia. Damage restricted to the semantic system (or the lexical-semantic route) does not yield surface dyslexia, because the preserved lexical-nonsemantic route activates correct phonology for exception words. Thus, the dual-route model interprets the commonly observed association between semantic impairment and surface dyslexia as non-causal, arising from damage to two independent (but probably adjacent) brain regions that underpin semantic representations and the lexical-nonsemantic route for reading.

The traditional view of reading aloud in Japanese, importing this classical dual-route conception, has typically assumed that kanji words and high-frequency kana words are processed by the lexical-nonsemantic route (and possibly also by the lexical-semantic route in the case of kanji words) (e.g. Hayashi, Ulatowska, & Sasanuma, 1985; Sasanuma, 1987). Low-frequency kana words and kana nonwords are supposed to be pronounced by a nonlexical route consisting of kana-mora correspondence rules. Damage to the lexical route explains a selective deficit in reading aloud kanji words, coupled with preserved ability to read aloud kana words and nonwords. Furthermore, because the efficiency of the lexical route is assumed to be frequency-modulated, its partial disruption produces a more severe deficit for low-frequency kanji words, as observed in TI's performance. The simple assumption that kanji words are processed by a word-specific lexical route, however, cannot account for the consistency effects observed in TI's performance, including his relatively preserved ability to read lower-frequency kanji words with consistent characters. As a result of a gradual awakening to the possibility of subword-level translation from orthography to phonology in kanji (Wydell, Butterworth, & Patterson, 1995), some researchers adopting a dual-route perspective have suggested that the nonlexical route applies to kanji characters (Saito, Masuda, & Kawakami, 1999; Tamaoka & Hatsuzuka, 1998), although the nature of this nonlexical translation has not been described in detail. According to this proposal, even in the absence of adequate lexical processing, a spared nonlexical route should provide the correct pronunciation for consistent and inconsistent-typical words, and the statistically typical (and therefore correct) pronunciation for nonwords.⁶ Again, however, this proposal cannot

account for graded effects of consistency in reading aloud words and nonwords—especially TI's poorer performance for inconsistent-typical than consistent words—because the concatenation of the most typical pronunciation for each component character constitutes the correct pronunciation for both consistent and inconsistent-typical words. Furthermore, if the nonlexical route provided the statistically most typical pronunciation for each component character, TI (as well as normal readers) should always assign the typical pronunciation to kanji nonwords. This is not the case: both TI and normal readers assign pronunciations to the component characters of nonwords in a manner graded by their degree of consistency.

Although there are several ways, in the dual-route framework, to account for the graded consistency effects observed in TI's performance, we describe briefly one candidate originally suggested in the dual-route cascaded (DRC) model (Coltheart et al., 1993; Coltheart et al., 2001). In a possible Japanese version of the DRC model, a word (e.g. 神経) activates not only its orthographic representation but also (though of course not as strongly) representations of words orthographically similar to the target (e.g. neighbors like 神父/siN-pu/, 神話/siN-wa/, 神様/kami-sama/). Also a nonword (e.g. 神間) weakly activates orthographic representations of real-word neighbors. These activations feed forward, in a cascaded manner, to their phonological representations and then to the phoneme system, where the outputs from the lexical-nonsemantic and the nonlexical routes are combined. This kind of cascaded processing may result in activation of nontypical (e.g. /kami/) pronunciations in inconsistent words and nonwords but not in consistent words and nonwords. If this kind of cascaded processing operates even in a partially damaged lexical-nonsemantic route, or if the nature of malfunction in this route in fact allows more neighborhood activation than in normal operation, then a graded consistency effect may emerge in surface dyslexia.

An alternative family of models able to predict and interpret graded effects of consistency observed in English-speaking readers is the triangle model instantiated in connectionist networks with several different specific architectures (e.g. Plaut et al., 1996; Seidenberg & McClelland, 1989). The triangle model assumes that a word is represented as a distributed pattern of activity over neuron-like processing units in separate orthographic, phonological, and semantic domains, each of which communicates with the other two domains via pools of hidden units. Activation in any one domain propagates automatically to both of the others. In reading aloud, the orthography of a given word activates its pronunciation primarily through connections from orthographic to phonological units. This computation is more efficient for high- than low-frequency words because the network's connection weights are gradually set during learning, when the system encounters high-frequency words many times. The computation is also most efficient for consistent words like GLOBE, intermediate for

⁶ Coltheart et al. (2001, p. 236) argued that the distinction between a lexical and nonlexical route cannot apply to Chinese or Japanese kanji because monosyllabic nonwords cannot be written in these scripts (The authors presumably mean single-character nonwords, and referred to 'monosyllabic' nonwords because one character always equals one syllable in Chinese. As illustrated earlier, one character of Japanese kanji may correspond to one, two or three morae). However, this description is insufficient because many kanji characters appear only as constituent characters of multiple-character kanji words and/or kanji-kana compound verbs or adjectives, but not as single-character kanji words. Therefore these characters, which might be thought of as a kind of bound morpheme, are single-character nonwords even if they have meaning. Furthermore, as described in the present study, in two-character kanji strings, the distinction between words and nonwords seems reasonable.

inconsistent-typical words like *STOVE*, and least efficient for inconsistent-atypical words like *GLOVE* because, in a distributed system, computation of the phonology of a target word is affected by neighboring words with a similar spelling pattern.

Plaut et al. (1996) implemented several different orthography-to-phonology networks with the goal of simulating frequency-by-regularity effects in normal readers and patterns of performance in surface dyslexia. In one simulation consisting merely of an orthography-to-phonology network, which successfully simulated frequency-by-regularity effects on normal response times, increasing levels of damage to this network eroded performance not only for exception words but also for regular words and nonwords. This result is incongruent with the fact that surface dyslexic patient may demonstrate preserved regular-word and nonword reading even when the impairment for exception words is profound (Funnell, 1996; McCarthy & Warrington, 1986). Another orthography-to-phonology network was therefore trained with additional input to phonology, intended to mimic phonological activation arising from semantics. Because the orthography-to-phonology network is inherently inefficient for lower-frequency words with less typical spelling-sound correspondences, the premise underlying this computational choice was that additional activation of phonology from semantics should be more critical for correct pronunciation of these “inefficient” words than for items that benefit from high word frequency or typical spelling-sound correspondences and are therefore processed efficiently just on the basis of orthography-to-phonology connections. When such putative semantic support was reduced in this version of the triangle model (Plaut et al., 1996), the network successfully simulated impaired reading of low-frequency exception words, coupled with preserved regular-word and nonword reading in surface dyslexic patients, as well as the neighborhood consistency effects observed in an English patient with profound semantic impairment and striking surface dyslexia (Patterson & Behrmann, 1997).

It is important to note that, because this version of the triangle model assumes a causal relationship between semantic impairment and surface dyslexia, its adequacy is challenged by any observations of semantically impaired patients whose reading does not reveal a surface dyslexic pattern. There have been several reports of this dissociation (Cipolotti & Warrington, 1995; Gerhand, 2001; Lambon Ralph et al., 1995), and we therefore acknowledge that this framework requires further development (see Plaut (1997) for an initial computational attack on this problem). It is nevertheless worth noting that a very small number of observations of a dissociation between substantial semantic impairment and preserved exception-word reading has to be evaluated in the context of by-now hundreds of observations of the association between semantic impairment and surface dyslexia in many different languages; and furthermore that—although traditional neuropsychological reasoning gives more weight

to one reported dissociation than to 100 observations of an association—this is not a universally accepted point of view. For example, “This is not to say that all instances of isolated strong dissociations are theoretically useless . . . but they must be approached with a degree of wariness, pending the demonstration of their high prevalence . . . [and they] should not be treated with unrestrained enthusiasm as the major tool of neuropsychological discovery and theory-building” (Goldberg, 1995, p. 195).

To return to models of reading aloud in Japanese: we propose that an interpretation of surface dyslexic reading as the manifestation of an intact orthography-to-phonology computation under the reduction of additional phonological activation from semantics is also useful to interpret and simulate surface dyslexia in Japanese. One enterprise along this line is a connectionist network for computing phonology from orthography of two-character kanji words developed by Ijuin et al. (1999). In this network, the orthography of a kanji word is represented as a grid or bitmap pattern over hundreds of visual units, each of which has an on or off value as used in computer fonts. The phonology of a word is represented as an activation pattern over an array of phoneme units, each of which represents a specific phoneme in a specific mora position. Thus, words (and nonwords) are represented over the same set of orthographic and phonological units shared across all items; there are no word-specific lexical nodes. The network has successfully simulated a frequency effect, a graded consistency effect, and the interaction between these two variables observed in skilled readers’ response times in reading aloud the same set of two-character kanji words employed in the present assessment of TI’s word-reading performance. It will therefore be of substantial theoretical interest to determine whether this network, when trained with additional input representing phonological activation arising from semantics, demonstrates surface dyslexic reading when this semantic support is reduced.

In summary, the graded consistency effects observed in TI’s reading aloud of kanji words and nonwords constitute important evidence for frequency and similarity-based orthography-to-phonology translation in Japanese kanji, and thus provide significant information to constrain the development of models of reading in Japanese and other nonalphabetic scripts.

Acknowledgements

We are grateful to Dr. Naruhiko Maki for taking responsibility for TI’s medical care while MI (the third author) was on study leave in Cambridge, UK.

References

- Amano S., & Kondo T. (1999). *NTT database series: Lexical properties of Japanese (I)*. Tokyo: Sansendo [in Japanese].

- Amano, S., & Kondo, T. (2000). *NTT database series: Lexical properties of Japanese (II)*. Tokyo: Sanseido [in Japanese].
- Behrmann, M., & Bub, D. (1992). Surface dyslexia and dysgraphia: dual routes, single lexicon. *Cognitive Neuropsychology*, 9, 209–251.
- Breedin, S. D., Saffran, E. M., & Coslett, H. B. (1994). Reversal of the concreteness effect in a patient with semantic dementia. *Cognitive Neuropsychology*, 11, 617–666.
- Chan, D., Fox, N. C., Scahill, R. I., Crum, W. R., Whitwell, J. L., Leschziner, G., Rossor, A. M., Stevens, J. M., Cipolotti, L., & Rossor, M. N. (2001). Patterns of temporal lobe atrophy in semantic dementia and Alzheimer's disease. *Annals Neurology*, 49, 433–442.
- Cipolotti, L., & Warrington, E. K. (1995). Semantic memory and reading abilities: A case report. *Journal of the International Neuropsychological Society*, 1, 104–110.
- Coltheart, M. (1985). Cognitive neuropsychology and study of reading. In M. I. Posner, O.S.M. Marin (Eds.), *Attention and performance XI* (pp. 3–37). Hillsdale, NJ: Erlbaum.
- Coltheart, M., Curtis, B., Atkins, P., & Haller, M. (1993). Models of reading aloud: Dual route and parallel distributed processing approaches. *Psychological Review*, 100, 589–608.
- Coltheart, M., Rastle, K., Perry, C., & Ziegler, J. (2001). DRC: A dual route cascaded model of visual word recognition and reading aloud. *Psychological Review*, 108, 204–256.
- Folstein, M. F., Folstein, S. E., & McHugh, P. R. (1975). "Mini-mental state". A practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research*, 12, 189–198.
- Funnell, E. (1996). Response biases in oral reading: an account of the co-occurrence of surface dyslexia and semantic dementia. *The Quarterly Journal of Experimental Psychology*, A49, 417–446.
- Fushimi, T., Ijuin, M., Patterson, K., & Tatsumi, I. F. (1999). Consistency, frequency, and lexicality effects in naming Japanese Kanji. *Journal of Experimental Psychology: Human Perception and Performance*, 25, 382–407.
- Galton, C. J., Patterson, K., Graham, K., Lambon Ralph, M. A., Williams, G., Antoun, N., Sahakian, B., & Hodges, J. R. (2001). Differing patterns of temporal atrophy in Alzheimer's disease and semantic dementia. *Neurology*, 57, 216–225.
- Gerhard, S. (2001). Routes to reading: a report of a non-semantic reader with equivalent performance on regular and exception words. *Neuropsychologia*, 39, 1437–1484.
- Goldberg, E. (1995). Rise and fall of modular orthodoxy. *Journal of Clinical and Experimental Neuropsychology*, 17, 193–208.
- Graham, K. S., Patterson, K., & Hodges, J. R. (1994). The relationship between comprehension and oral reading in progressive fluent aphasia. *Neuropsychologia*, 32, 299–316.
- Graham, K., Patterson, K., & Hodges, J. R. (1995). Progressive pure anomia: Insufficient activation of phonology by meaning. *Neurocase*, 1, 25–38.
- Graham, N. L., Patterson, K., & Hodges, J. R. (2000). The impact of semantic memory impairment on spelling: evidence from semantic dementia. *Neuropsychologia*, 38, 143–163.
- Hayashi, M., Ulatowska, H., & Sasanuma, S. (1985). Subcortical aphasia with deep dyslexia: a case study of a Japanese patient. *Brain and Language*, 25, 293–313.
- Hodges, J. R., Patterson, K., Oxbury, S., & Funnell, E. (1992). Semantic dementia. Progressive fluent aphasia with temporal lobe atrophy. *Brain*, 115, 1783–1806.
- Hughes, C. P., Berg, L., Danziger, W. L., Coben, L. A., & Martin, R. L. (1982). A new clinical scale for the staging of dementia. *British Journal of Psychiatry*, 140, 566–572.
- Ijuin, M., Fushimi, T., Patterson, K., & Tatsumi, I. (1999). A connectionist approach to Japanese kanji word naming. *Psychologia: An International Journal of Psychology in the Orient*, 42, 267–280.
- Ito, K., Nakagawa, Y., Ikeda, M., Yamada, N., Hashimoto, M., & Tanabe, H. (1994). Category-specific word meaning impairments in Gogi aphasics. *Higher Brain Function Research*, 14, 221–229 [in Japanese with an English abstract].
- Jared, D., McRae, K., & Seidenberg, M. S. (1990). The basis of consistency effects in word naming. *Journal of Memory and Language*, 29, 687–715.
- Lambon Ralph, M. A., Cipolotti, L., & Patterson, K. (1999). Oral naming and oral reading: Do they speak the same language? *Cognitive Neuropsychology*, 16, 157–169.
- Lambon Ralph, M. A., Ellis, A. W., & Franklin, S. (1995). Semantic loss without surface dyslexia. *Neurocase*, 1, 363–369.
- Lambon Ralph, M. A., McClelland, J. L., Patterson, K., Galton, C. J., & Hodges, J. R. (2001). No right to speak? The relationship between object naming and semantic impairment: Neuropsychological evidence and a computational model. *Journal of Cognitive Neuroscience*, 13, 341–356.
- McCarthy, R., & Warrington, E. K. (1986). Phonological reading: phenomena and paradoxes. *Cortex*, 22, 359–380.
- Nakamura, H., Nakanishi, M., Hamanaka, T., Nakaaki, S., & Yoshida, S. (2000). Surface dyslexia and visual word recognition in Japanese. *Higher Brain Function Research*, 20, 136–144 [in Japanese with an English abstract].
- National Language Research Institute (1970). *A computational assessment of vocabulary in newspapers*. Report No. 37. Tokyo: Shuei Shuppan [in Japanese].
- National Language Research Institute (1976). *A study of uses of Kanji characters in modern newspapers*. Report No. 56. Tokyo: Shuei Shuppan [in Japanese].
- Nishio, M., Iwabuchi, E., & Mizutani, S. (1986). *Iwanami Japanese Dictionary*. 4th ed. Tokyo: Iwanami shoten [in Japanese].
- Papagno, C., & Capitani, E. (2001). Slowly progressive aphasia: a 4-year follow-up study. *Neuropsychologia*, 39, 678–686.
- Patterson, K., & Behrmann, M. (1997). Frequency and consistency effects in a pure surface dyslexic patient. *Journal of Experimental Psychology: Human Perception and Performance*, 23, 1217–1231.
- Patterson, K., & Hodges, J. R. (1992). Deterioration of word meaning: Implications for reading. *Neuropsychologia*, 30, 1025–1040.
- Patterson, K., Suzuki, T., Wydell, T., & Sasanuma, S. (1995). Progressive aphasia and surface alexia in Japanese. *Neurocase*, 1, 155–165.
- Plaut, D. C. (1997). Structure and function in the lexical system: Insights from distributed models of word reading and lexical decision. *Language and Cognitive processes*, 12, 767–808.
- Plaut, D. C., McClelland, J. L., Seidenberg, M. S., & Patterson, K. (1996). Understanding normal and impaired word reading: Computational principles in quasi-regular domains. *Psychological Review*, 103, 56–115.
- Rastle, K., & Coltheart, M. (1999). Serial strategic effects in reading aloud. *Journal of Experimental Psychology: Human Perception and Performance*, 25, 482–503.
- Raven, J. C. (1965). *Guide to using the Coloured Progressive Matrices*. London: HK Lewis.
- Rey, A. (1941). L'examen psychologique dans les cas d'encephalopathie traumatique. *Archives de Psychologie*, 28, 286–340.
- Rosen, W. G., Mohs, R. C., & Davis, K. L. (1984). A new rating scale for Alzheimer's disease. *American Journal of Psychiatry*, 141, 1356–1364.
- Saito, H., Masuda, H., & Kawakami, M. (1999). Subword activation in reading Japanese single kanji character words. *Brain and Language*, 68, 75–81.
- Sasanuma, S. (1980). Acquired dyslexia in Japanese: clinical features and underlying mechanisms. In M. Coltheart, K.E. Patterson, & J.C. Marshall (Eds.), *Deep dyslexia* (pp. 48–90). London: Routledge & Kegan Paul.
- Sasanuma, S. (1985). Surface dyslexia and dysgraphia: How are they manifested in Japanese. In: K. Patterson, R. McCarthy, C. Marshall, & M. Coltheart (Eds.), *Surface dyslexia* (pp. 225–249). Hillsdale, NJ: Erlbaum.
- Sasanuma, S. (1987). Reading disorder resulting from brain damage. In K. Goryo (Ed.), *Reading* (pp. 175–226). Tokyo: Tokyo Daigaku Shuppan-kai [in Japanese].

- Seidenberg, M. S., & McClelland, J. L. (1989). A distributed, developmental model of word recognition and naming. *Psychological Review*, 96, 523–568.
- Seidenberg, M. S., Plaut, D. C., Petersen, A. S., McClelland, J., & McRae, K. (1994). Nonword pronunciation and models of word recognition. *Journal of Experimental Psychology: Human Perception and Performance*, 20, 1177–1196.
- Shallice, T., Warrington, E., & McCarthy, R. (1983). Reading without semantics. *Quarterly Journal of Experimental Psychology*, 35A, 111–118.
- Shinagawa, F., Kobayashi, S., Fujita, K., & Maekawa, H. (1990). *Japanese Wechsler Adult Intelligence scale-revised*. Tokyo: Nihon Bunka Kagaku-sha [in Japanese].
- SLTA Committee (1997). *Standard Language Test of Aphasia Manual*. Tokyo: Shinkou Igaku Shuppan-sha [in Japanese].
- Snowden, J. S., Goulding, P. J., & Neary, D. (1989). Semantic dementia: a form of circumscribed cerebral atrophy. *Behavioural Neurology*, 2, 167–182.
- Tamaoka, K., & Hatsuzuka, M. (1998). The effects of morphological semantics on the processing of Japanese two-kanji compound words. *Reading and Writing: An Interdisciplinary Journal*, 10, 293–322.
- Tanabe, H., Ikeda, M., Nakagawa, Y., Yamamoto, H., Ikejiri, Y., Kazui, H., Hashikawa, K., & Harada, K. (1992). Gogi (word meaning) aphasia and semantic memory for words. *Higher Brain Function Research*, 12, 153–167 [in Japanese with an English abstract].
- Tanabe, H., Nakagawa, Y., Ikeda, M., Hashimoto, M., Yanada, N., Kazui, H., Nishikawa, T., & Okuda, J. (1996). Selective loss of semantic memory for words. In K. Ishikawa, J. L. McGaugh, & H. Sakata (Eds.), *Brain Processes and Memory* (pp. 141–152). Amsterdam: Elsevier Science.
- Treiman, R., Mullenix, J., Bijeljac-Babic, R., & Richmond-Welty, E. D. (1995). The special role of rimes in the description, use, and acquisition of English orthography. *Journal of Experimental Psychology: General*, 124, 107–136.
- 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.
- Wechsler, D. (1981). *WAIS-R manual*. New York: The Psychological Corporation.
- Wydell, T. N., Butterworth, B., & Patterson, K. (1995). The inconsistency of consistency effects in reading: The case of Japanese Kanji. *Journal of Experimental Psychology: Learning Memory and Cognition*, 21, 1155–1168.