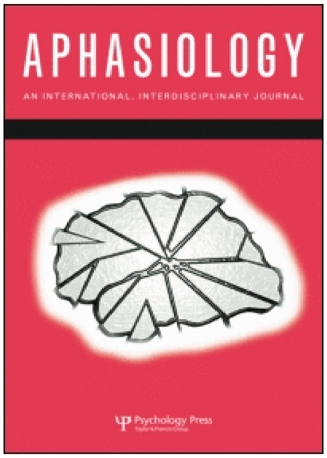

Aphasiology
Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t713393920

The effect of typicality on online category verification of inanimate category exemplars in aphasia

Online Publication Date: 01 September 2007
To cite this Article: Kiran, Swathi, Ntourou, Katerina and Eubank, Megan (2007) 'The effect of typicality on online category verification of inanimate category exemplars in aphasia', Aphasiology, 21:9, 844 - 866
To link to this article: DOI: 10.1080/02687030600743564
URL: http://dx.doi.org/10.1080/02687030600743564

APHASIOLOGY, 2007, 21 (9), 844–866

Psychology Press
Taylor & Francis Group

The effect of typicality on online category verification of inanimate category exemplars in aphasia

Swathi Kiran, Katerina Ntourou and Megan Eubank

University of Texas at Austin, TX, USA

Background: A previous study (Kiran & Thompson, 2003a) investigated the effect of typicality on online category verification of animate categories in patients with fluent or nonfluent aphasia and their normal controls. Results revealed a robust effect of typicality: typical examples were faster and more accurate than atypical examples of animate categories. Patients with fluent aphasia did not demonstrate the expected effects of typicality.
Aims: The aim of the present study was to extend this work to examine the effect of typicality on inanimate categories such as *furniture*, *clothing*, and *weapons*.
Methods & Procedures: Normal young, older, and aphasic individuals participated in an online category verification task where primes were superordinate category labels whereas targets were either typical or atypical examples of inanimate categories (e.g., *clothing*, *furniture*, *weapons*) or nonmembers belonging to animate categories. Aphasic participants were divided into two groups, semantic impairment group (SI) and no semantic impairment group (NSI), based on their performance on offline standardised semantic processing tests. The reaction time to judge whether the target belonged to the preceding category label was measured.
Outcomes & Results: Results indicated that all four groups were significantly faster and more accurate on typical examples compared to atypical examples. Further, differences emerged in the processing of categories, wherein responses to *clothing* were more accurate than responses to *furniture* or *weapons*. In the SI group, representation of typical examples and atypical examples were impaired, as evidenced by poor accuracy rates.
Conclusions: The present experiment demonstrated the typicality effect in normal individuals and in individuals with aphasia. Further, differences emerged in the processing of categories, where responses to *clothing* were more accurate than responses to *furniture* or *weapons*.

TYPICALITY IN INANIMATE CATEGORIES IN APHASIA

Much research in psychology has focused on the representation of semantic categories. The classical view of categories being represented by a set of defining features that allows equivalent probability of membership for all members (Bruner, Goodnow, & Austin, 1956) has been replaced by the observation that not all

Address correspondence to: Swathi Kiran PhD, CMA 7.206, Department of Communication Sciences & Disorders, University of Texas at Austin, Austin, TX, 78712, USA. E-mail: s-kiran@mail.utexas.edu

This research was supported by NIDCD # DC006359-03 to the first author. The authors wish to thank Shilpa Shamapant and Lisa Edmonds for their assistance during the development stages of the project.

http://www.psypress.com/aphasiology DOI: 10.1080/02687030600743564

members of a category are equal (Posner & Keele, 1968; Rosch, 1973, 1975). This idea of "graded membership" within categories has been supported by several studies that have shown differences in lexical processing between typical and atypical exemplars of a category, with typical examples receiving preferential processing, and this phenomenon has been labelled the typicality effect (Hampton, 1979; McCloskey & Glucksberg, 1978; Posner & Keele, 1968; Rosch, 1973, 1975).

Representation of typicality

The typicality effect has been shown using various experimental paradigms including (a) participants' ratings of typicality of items within a category (Rosch, 1975; Rosch & Mervis, 1975; Uyeda & Mandler, 1980), (b) the order in which category items are learned (Posner & Keele, 1968; Rosch, 1973; Rosch & Mervis, 1975), (c) probability of item output within a category (Mervis, Catlin, & Rosch, 1976; Rosch, 1975; Uyeda & Mandler, 1980), (d) expectations generated by category names (Rosch, 1975), and (e) category naming frequency (Casey, 1992; Hampton, 1995). More relevant to the present experiment, typicality predicts verification time for category membership (Hampton, 1979; Larochelle & Pineau, 1994; McCloskey & Glucksberg, 1978; Rips, Shoben, & Smith, 1973; Smith, Shoben, & Rips, 1975). All these studies have found faster reaction times for typical examples than for atypical examples during a category verification task. The results of the online category verification tasks can be explained by the spreading activation theory (Collins & Loftus, 1975), which suggests that concepts are represented as nodes within a semantic network. When a concept is processed, the appropriate node is activated and this activation automatically spreads along the connections of the network, gradually decreasing in strength based on semantic relatedness.

Although typicality of examples is a measure of the inherent representativeness of each example within a category, it is argued that this variable can often be confounded by familiarity and frequency of specific examples. McCloskey (1980) showed that familiarity of examples of a category correlated with rated typicality, and once this factor was partialled out the effect of typicality on reaction time was reduced but not eliminated. Likewise, Ashcraft (1978) reported that people are more familiar with typical than atypical examples of categories, and thus can access them more readily. However, Mervis and colleagues (1976) found no correlation between how often an example is produced in response to the category name and word frequency. Other studies have shown that typicality ratings obtained for many categories are relatively unrelated to certain properties of the examples such as familiarity and frequency (Mervis et al., 1976; Rosch, 1975). Finally, a study by Malt and Smith (1982) showed that variations in typicality do not always coexist with variations in familiarity of the category examples, although familiarity may play a role depending on the nature of the category. For example, for inanimate categories such as *clothing*, in which most examples are well known, semantic features such as physical or functional characteristics of the examples may be the most important determinant of typicality ratings. However, for other categories such as *trees*, familiarity may be a more important determinant. In summary, the extent to which familiarity influences typicality varies according to the different categories.

There are several models proposed to explain the typicality effect (see Komatsu, 1992, for a review). Of these, the prototype model is consistent with the theoretical premise of the present work and hence will be described in some detail. According to this model, a category prototype is a generic representation of the common features of the category taken as a whole. Therefore, across categories, there is a set of features that exert differential weights in the definition of a prototype (Hampton, 1995). The explicit versions of the prototype model (Hampton, 1993; Rosch & Mervis, 1975) propose that similarity to the prototype increases with the number of overlapping features, and hence typical examples have more overlap with the prototype than atypical examples. Further, Rosch and Mervis (1975) found that the degree to which a given member possessed attributes in common with other members was highly correlated with the degree to which it was rated typical of the category, i.e., typical members (e.g., *robin*) shared more features with other birds (e.g., *wren*, *finch*), whereas, atypical members (e.g., *ostrich*) shared fewer features with examples of birds. Consequently, similarity judgements for a category place typical examples in the centre of a multidimensional semantic space (and the prototype) and atypical examples furthest away from the prototype (Rips et al., 1973; Rosch & Mervis, 1975).

Organisation of animate/inanimate categories

To further understand the mechanisms underlying typicality within a category, it is important to consider exemplar typicality within the broader framework of representation and structure of semantic categories in general. An important organising principle of semantic categories is their representation in terms of a distributed set of correlated features (Devlin et al., 2002; McRae, de Sa, & Seidenberg, 1997; Tyler & Moss, 2001; Tyler, Moss, Durrant-Peatfield, & Levy, 2000). Within a category, individual items vary in their similarity to other members of the category, and therefore, in their distance from the centre of the semantic cluster (Tyler & Moss, 2001). Much research has focused on the distinction between the representation of animate and inanimate categories. In general, studies have shown that perceptual features are more important than functional features in category membership decisions of animate categories, whereas functional features are more important than perceptual features in category membership of inanimate categories (Barton & Komatsu, 1989; Devlin et al., 2002; Diesendruck & Gelman, 1999; Keil, 1989; Vanoverberghe & Storms, 2003). Supportive empirical evidence also comes from McRae et al., (1997) who found that animate categories have more intercorrelated features than inanimate categories.

Typicality and the animate/inanimate distinction

Typicality of examples influences the differential representation of animate and inanimate categories in two respects. First, representation of features within a category is partially guided by typicality of examples. For instance, Garrard, Lambon Ralph, Hodges, and Patterson (2001) developed typicality ratings and a database of semantic features for specific items across several categories. Generated features were then classified as sensory (e.g., *a duck has webbed feet*), functional (e.g., *a duck can fly*), or encyclopaedic (e.g., *a duck is found near water*), and analysed with regard to their dominance (frequency of elicitation for a given item) and distinctiveness (the percentage of category members for which the feature was characteristic). Results showed that a greater proportion of shared features were observed among typical items in animate categories than in inanimate categories. Further, animate categories were associated with a higher ratio of sensory to functional features and a higher intercorrelation between features than inanimate categories (Garrard et al., 2001).

Likewise, Vanoverberghe and Storms (2003) examined the importance of feature type in animate and inanimate categories in two tasks, feature generation and typicality ratings. The authors observed that inanimate categories have more functional features, while both animate and inanimate categories consist of perceptual features. Additionally, in prediction for typicality ratings, perceptual features were more important for animate categories whereas functional features were more important for inanimate categories.

More relevant to the present study, a second situation where the interaction between typicality and the animate/inanimate category distinction comes into play is during category membership decisions. For instance, according to Diesendruck and Gelman (1999), typicality was more relevant for inanimate categories than for animate categories. The authors observed that participants tended to make absolute category membership judgements of animate categories such as *animals* whereas judgement of inanimate categories was more graded in terms of typicality. Barr and Caplan (1987) reported similar findings from typicality and category membership ratings from 13 different categories (animate and inanimate). Categories such as *animals* were given more absolute judgements of category membership than categories such as *artefacts* (Barr & Caplan, 1987). Likewise, Estes (2003) found that inanimate categories were more graded than animate categories (e.g., *rug* was more likely to be judged a partial member of *furniture* than *tomato* was judged a partial member of *fruit*).

Typicality in aphasia

While typicality has received a great deal of attention in normal psycholinguistic experiments, few studies have investigated the effect of typicality of category exemplars in aphasia (Grober, Perecman, Kellar, & Brown, 1980; Grossman, 1981; Hough, 1993; Kiran & Thompson, 2003a). In an off-line experiment, Hough (1993) examined category verification and exemplar generation in fluent and nonfluent aphasia using common and goal-derived categories. Results revealed that patients with fluent and nonfluent aphasia exhibit difficulty in accessing peripheral category examples, indicating impoverished representations at the boundaries of categories. Likewise, during an exemplar generation task, Grossman (1981) found that nonfluent aphasic individuals named significantly more members from the central field of a superordinate category than less-central instances. On the other hand, fluent aphasic individuals often violated the category boundary by naming examples that did not belong to the category, and produced fewer central field members. Grober et al. (1980) compared the performance of anterior, posterior aphasic individuals and control participants on a category judgement task. Both anterior and posterior aphasic individuals judged membership of typical exemplars and semantically unrelated nonmembers (items situated far from the category boundaries) with high accuracy, but their performance diverged with atypical members. Accuracy on judgement tasks dropped from 100% for typical members to 85% for atypical members for the anterior aphasic patients, while for the posterior aphasic patients, accuracy in judgement tasks dropped from 95% for typical members to 66% for atypical members.

In a recent study that investigated the typicality effect on category verification in individuals with aphasia (Kiran & Thompson, 2003a), participants were divided in four distinct groups: normal young, normal older, Broca's and Wernicke's individuals with aphasia. An online category verification task was conducted using animate categories (*birds*, *vegetable*, *fish*) as primes and typical, atypical, and nonmember items as targets. Results indicated that young, older, and Broca's individuals demonstrated faster reaction times for typical than atypical examples. Contrastingly, Wernicke's patients did not show preferential processing for typical versus atypical examples, demonstrated by statistically insignificant difference in reaction times between typical and atypical instances of the animate categories. Overall, all four groups made more errors on atypical than typical exemplars, with the aphasic individuals, and especially Wernicke's patients, demonstrating greater number of errors than normal controls. This finding suggests that items that are at the boundaries of a category (atypical items) are more prone to error than the typical and nonmember examples, and this effect is magnified in participants with aphasia.

The present study

The present study sought to extend the findings of Kiran and Thompson (2003a) to further understand the effect of typicality in aphasia. These two projects comprised part of a broader effort to explain typicality as a variable for rehabilitation of naming deficits in aphasia. Our previous work (Kiran & Thompson, 2003b) has demonstrated that training atypical examples facilitates generalisation to untrained typical examples but not vice versa. These findings have led us to suggest that, within a category, atypical examples are more complex than typical examples (Kiran, in press). One line of evidence contributing to this hypothesis is the longer reaction times for atypical examples compared to typical examples, indicating that atypical examples are represented further away from the category prototype and typical examples in terms of time and space (for a similar proposal equating processing time with complexity see Gennari & Poeppel, 2003).

Given the differences in the degree of gradation observed between processing of animate and inanimate categories, our previous study examined animate categories, whereas the present study was focused on inanimate categories. The present study also differed from the previous study in the assignment of the patient population. Instead of assigning aphasic participants into diagnostic classificatory groups (such as anterior/posterior, fluent/nonfluent, or Broca's/Wernicke's), participants in the present study were divided into two groups based on their performance on subtests of semantic processing as assessed by the Psycholinguistic Assessment of Language Processing Abilities in Aphasia (Kay, Lesser, & Coltheart, 1992). Although this distinction deviates from the traditional approach to examine semantic processing in aphasia, three lines of evidence justify such a classification. Work by Basso and colleagues (Basso, Lecours, Moraschini, & Vanier, 1985)—who studied the CT scans of 267 aphasic individuals and tested their oral expression, auditory verbal comprehension, and repetition skills—has shown that aphasic patients cannot be categorised into different language syndromes (e.g., Broca's, Wernicke's, conduction, etc.) solely based on their site of lesion. Further, recent reviews of neuroimaging studies suggest that semantic processing involves a network of activation that may include both the posterior and anterior regions (for reviews see Cabeza & Nyberg, 2000; Thompson-Schill, 2003), questioning the validity of assigning participants into groups based on their aphasia type in order to assess semantic processing abilities. Finally, there is evidence suggesting that nonfluent Broca's aphasic patients do not always perform normally on semantic processing tasks (Del Toro, 2000; Milberg, Blumstein, & Dworetzky, 1987) and it is possible that fluent anomic patients do not always demonstrate semantic processing impairments. Finally, an important goal of this study was to understand differences in online semantic processing between two groups relative to offline behavioural performance on related semantic processing tasks. Certainly, such an assignment is consistent with our treatment studies (e.g. Kiran & Thompson, 2003b) where participants are inducted into treatment based on whether or not they present with semantic impairments.

Given that the aim of the study was to provide further proof for the typicality effect, it was predicted that all four participant groups will demonstrate faster reaction times and fewer errors for typical compared to atypical examples. Unlike animate categories, however, we expected representation of inanimate categories to be more graded i.e., typicality effects were expected to be larger and more robust than animate categories. In addition, we expected the no semantic impairment group (NSI) and semantic impaired group (SI) to perform differently from their normal controls on these two measures of typicality (latency and errors). We predicted similar performance by participants across the three inanimate categories (*furniture*, *clothing*, *weapons*). Finally, we were also interested in validating patient factors that were most predictive of performance on the reaction time tasks, so that future studies could incorporate classification of patient groups based on performance relevant to semantic priming than on aphasia type.

METHOD

Participants

A total of 10 normal young (M=27 years, age range=21–39 years), 10 normal older (M=59 years, age range=41–82 years), and 17 aphasic individuals (M=64.4 years, age range=47–77 years) participated in the experiment. The young and older participants were recruited from the University of Texas at Austin. All these participants had normal or corrected to normal vision, normal hearing, and had at least a high-school education. Exclusionary criteria included neurological disorders such as stroke, transient ischaemic attacks, Parkinson's disease, Alzheimer's disease, psychological illnesses, history of alcoholism, learning disability, seizures, and attention deficit disorders. Handedness was not controlled in these individuals.

The 17 aphasic participants were selected from the Aphasia Research Laboratory subject pool initially recruited from local area-wide hospitals. Several participant selection criteria were met in order for them to be included in the study: (a) diagnosis by a neurologist of a stroke in the left hemisphere (encompassing the grey and or white matter in and around the perisylvian area confirmed by a CT or MRI scan), (b) onset of stroke at least 9 months prior to participation in the study, (c) no concomitant visual or cognitive deficits as determined by a certified speech language pathologist, (d) at least a high-school diploma, and (e) native speaker of English. Except for one patient (P3), all participants were right-handed. Nine participants presented with right-sided paralysis concurrent with their aphasia.

Pretesting language tests were administered to all the participants to ensure diagnosis of aphasia as measured by calculation of Aphasia Quotient (AQ) of the *Western Aphasia Battery* (*WAB*) (Kertesz, 1982). Naming performance was assessed through the *Boston Naming Test* – second edition (Kaplan, Goodglass, & Weintraub, 2001) and the naming subtest of the *WAB*. Semantic processing was assessed using selected subtests from the *Psycholinguistic Assessment of Language Processing in Aphasia* (*PALPA*; Kay et al., 1992): spoken word to picture matching, written word to picture matching, auditory synonym judgement, and written synonym judgement. Performance on only the latter two tests was used to divide participants either into a semantic impairment group (SI group) or a no semantic impairment group (NSI group) for the following reasons. First, it was hypothesised that word to picture matching tasks were inherently easier than the experimental task that required verification of written word pairs. As expected, most participants performed with relatively high accuracy on this task since the presentation of picture stimuli permits easier access to information. Incidentally, the spoken word to picture matching task has been criticised for a number of stimuli confounds that affect its interpretability (Cole-Virtue & Nickels, 2004). Instead, for the present study, average performance on the auditory and synonym judgement task was selected to divide participants into two groups. The synonym judgement task was deemed to be most similar to the experimental tasks in terms of task processing requirements; participants were required to encode each word, retrieve the meaning, and judge similarity. Participants with an average performance above 85% accuracy across the auditory and visual modality were assigned to the NSI group. Alternatively, participants with an average performance of less than 85% accuracy were assigned to the SI group. Participant 6 obtained an average of 84.9% accuracy and was assigned to the NSI group. Individual information regarding performance by participants is shown in Table 1.

The NSI group consisted of eight participants (Mean=64.75 years, range=53–72 years) matched in age to the SI group which consisted of nine participants (Mean=63.3 years, range=47–77 years). The mean AQ of the participants of the NSI group was 83.6 with a range of 62.7–93.8, whereas the mean AQ of the participants of the SI group was 69.8 with a range of 46.4–79.5. For the NSI group, the mean score for auditory comprehension, as assessed by the *WAB,* was 95% accuracy (range=83–100%), and for the SI group it was 83% accuracy (range=65–96%). The mean score on the naming subtest of the *WAB* for the NSI group was 88% accuracy (range=69–100%) and for the SI group the mean score was 71% accuracy (range=46–82%). The mean score on the *BNT* for the NSI group was 73% accuracy (range=22–96%) and for the SI group the mean score was 42% accuracy (range=13–68%).

On the spoken word to picture matching tasks within the *PALPA*, the NSI group (M=92%, range=85–95%) was better than SI aphasic participants (M=98%, range=95–100% accuracy). Similarly, on the written word to picture matching task, the NSI group was superior (M=99%, range=97–100% accuracy) to the SI group (M=94%, range=92–100% accuracy). The NSI group was also more accurate at judging auditory word pairs as synonyms (M=93%, range=99–98% accuracy) than the SI group (M=73%, range=66–83%). Finally, on the written word pair synonym judgement task, the NSI group was superior (M=91%, range=81–98%, accuracy) to the SI group (M=77%, range=66–83%, accuracy). However, both groups evinced visual lexical decision abilities within normal limits for real words and nonwords

TABLE 1
Participant information

Participant	*Group*	*Age*	*WAB AQ*	*WAB Dx*	*WAB-naming*	*WAB-AC*	*BNT*	*PALPA-SWPM*	*PALPA-WWPM*	*PALPA-ASJ*	*PALPA-WSJ*
P1	NSI	72	93.8	Anomic	92%	95%	82%	100%	100%	98%	98%
P2	NSI	60	91	Anomic	91%	98%	96%	100%	100%	95%	93%
P3	NSI	73	93.4	Anomic	87%	100%	76%	95%	100%	95%	86%
P4	NSI	71	84.6	Conduction	95%	100%	66%	100%	100%	88%	93%
P5	NSI	63	79.5	TM	90%	94%	88%	100%	100%	93%	95%
P6	NSI	64	71.8	Broca	85%	94%	58%	97%	97%	88%	81%
P7	NSI	53	92.1	Anomic	100%	97%	96%	100%	100%	95%	98%
P8	NSI	62	62.7	Broca	69%	83%	22%	95%	100%	91%	90%
AVERAGE		*64.75*	*83.61*		*88.63%*	*95.19%*	*73%*	*98.44%*	*99.69%*	*93.06%*	*91.98%*
P9	SI	56	56.7	Conduction	51%	89%	26%	92%	95%	78%	81%
P10	SI	77	73.4	TM	82%	89%	50%	87%	90%	66%	81%
P11	SI	67	77.8	Anomic	78%	94%	26%	97%	97%	83%	70%
P12	SI	77	72.5	Conduction	78%	78%	16%	95%	92%	68%	66%
P13	SI	60	80.6	Anomic	69%	80%	68%	97%	100%	78%	80%
P14	SI	47	68.5	Broca	80%	78%	60%	97%	95%	76%	83%
P15	SI	54	46.4	Broca	46%	65%	13%	85%	97%	66%	75%
P16	SI	60	78.4	Broca	82%	96%	71%	87%	92%	66%	83%
P17	SI	72	74.6	TM	78%	83%	50%	95%	95%	81%	76%
AVERAGE		*63.33*	*69.87*		*71.56%*	*83.72%*	*42.22%*	*92.78%*	*94.89%*	*73.96%*	*77.52%*

Individual information and performance on WAB (Kertesz, 1982), BNT (Kaplan et al., 2001), and selected PALPA (Kay et al., 1992) subtests for 17 aphasic participants. AC=auditory comprehension; SWPM=spoken word–picture matching; WWPM=written word–picture matching; ASJ=auditory synonym judgement; WSJ=written synonym judgement; TM=transcortical motor aphasia.

(NSI group M=98%, range=92–100%, SI group M=96%, range=92–100%). All participants were able to read and comprehend single words, which were tested through the reading subtest of the WAB (NSI group M=85%, range=72–97%; SI group M=70.4%, range=53–80%).

A specific aim of the study was to select two groups that were comparable on most aspects of language except their semantic processing abilities. At the same time, in order to ensure that all participants were able to complete the experimental task, participants with gross semantic impairments such as those evident on the picture word matching tasks were specifically not recruited, as they would be unable to complete the experimental task with reasonable accuracy. Therefore both groups had some degree of overlap in scores although there were differences in performance, the influence of which was addressed by a regression analysis. For instance, to eliminate the possibility that participants in the SI group had a global impairment that encompassed semantic processing abilities, the difference in WAB AQ between the two groups was addressed as a part of a regression analysis.

Stimuli

Three inanimate categories (*furniture*, *clothing*, *weapons*) and their examples as well as nonmembers were used in the present experiment. For the selection of the category and stimuli, which was part of another study (Kiran, 2002), 20 normal young and older adults were asked to provide as many instances of 12 categories that included *furniture*, *clothing*, *weapons*. Then a different group of an equal number of normal young and older adults was presented with each item generated by the first group, and was asked to judge typicality of each example within the category. The rating was based on a 7-point scale that has been used in other studies (Diesendruck & Gelman, 1999; Rosch, 1975; Uyeda & Mandler, 1980). A rating of 1 stood for a very good example of the category, 4 corresponded to a moderate example, and a rating of 7 was given to very poor examples. To avoid the influence of familiarity on the ratings, participants were asked to mark *U* for unfamiliar items (Malt & Smith, 1982). Once the ratings were completed, average rating score, standard deviation, and z scores for each example of each category were calculated across the 20 raters. The three experimental categories (*furniture*, *clothing*, *weapons)* were chosen based on several inclusionary and exclusionary criteria applied to stimuli (see Kiran, 2002) for details. Results obtained are consistent with other published typicality ratings (Rosch, 1975; Uyeda & Mandler, 1980).

The stimuli for each of the three inanimate categories were selected based on the z scores of the average typicality ratings for every rated item. For the category *weapons* the typical z scores ranged from -1.3 to -0.38 and the atypical z scores ranged from 1.16 to .08. For the category *furniture,* these values were -1.37 to -0.42 (typical) and 1.12 to 0.41 (atypical). For *clothing,* the z scores ranged from -1.22 to -0.44 (typical) and from -0.01 to 0.05 (atypical). In addition to the 15 typical and 15 atypical examples for each category, 45 typical members from three animate categories *fruit*, *vegetables*, and *animals* (N=15 each) were selected as nonmembers. These stimuli were chosen to ensure nonmembership in the test categories.

In order to rule out the possibility that typicality ratings correlated with written word frequency of the stimuli, a single 3 (typicality: typical, atypical, nonmembers) $\times$ 3 (category: *clothing*, *furniture*, *weapons*) ANOVA was performed on the written word frequency (Frances & Kucera, 1982). Results revealed no significant effects for category, $F(2, 126)=2.9$, $p=.06$, or typicality, $F(2, 126)=0.11$, $p=.89$, or interaction between category and typicality, $F(4, 126)=0.12$, $p=.97$. Another factor that could possibly confound the typicality ratings is familiarity of examples (Ashcraft, 1978; McCloskey, 1980). Another 3 (typicality: typical, atypical, nonmembers) $\times$ 3 (category: *clothing*, *furniture*, *weapons*) ANOVA was performed on familiarity ratings obtained from the MRC Psycholinguistic Database (Wilson, 1988). Results revealed no significant effects for category, $F(2, 74)=2.1$, $p=.12$, or typicality, $F(2, 74)=0.10$, $p=.89$. Also, no interaction effects were found for category and typicality, $F(4, 74)=2.18$, $p=.07$. See Appendix for a list of stimuli used in the experiment.

In sum, each experimental inanimate category (*furniture*, *clothing*, *weapons*) contained 15 typical, 15 atypical, and 15 nonmembers respectively, resulting in a total of 135 items ($15 \times 9=135$). Each of these items was paired with a superordinate category label during online presentation. Three blocks of 45 word pairs were constructed. Within each block the presentation of stimuli were randomised. The entire experiment took approximately 25 minutes and each block took an average of 8 minutes.

Procedure

Participants were seated in front of a computer with their non-dominant hand placed on the keyboard. In order to control for differences in the nine participants who presented with right-sided paralysis, participants from all four groups were required to perform the task with their nondominant hand. A Dell Latitude D600 (PC) laptop computer loaded with Superlab Pro (Cedrus Corporation, Phoenix Arizona) was used to present stimuli and record responses (reaction time and errors). The participants were presented with written instructions on the computer screen followed by verbal clarifications about the task. They were told that they would see a superordinate category label followed by a word item, and had to decide if that word belonged to the preceding category label as accurately and quickly as possible. The experiment began with a practice using a different set of stimuli during which no data were collected. Participants performed the practice task until accuracy on the task was 90%, with feedback provided regarding accuracy after each stimulus pair. Approximately 10 pairs were viewed by each participant.

During the experiment, the superordinate category label was flashed in the centre of the screen for 1000 ms, in red letters (font size 36). Interstimulus interval (ISI) between the presentation of the prime and target was 200 ms (Rosch, 1975). The target item was presented in black letters (font size 36) at the centre of the screen and remained on the screen until the participant made the category verification decision by pressing one of the two designated keys on the computer keyboard to signal a "yes" or "no" response. Finally, the intertrial interval (ITI) was 1500 ms. Participants were given the option of a 20-second break between Block 1–Block 2 and between Block 2–Block 3.

Data analysis

Mean reaction times (RT) and standard deviations, as well as accuracy rates (percentage of errors) for the typical, atypical, and nonmembers for the three inanimate categories (*furniture*, *clothing*, and *weapons*) were calculated separately for the four groups. The independent variables in the experiment were group (young, older, SI, and NSI groups), category (*furniture*, *clothing*, *weapons*), and typicality *(typical*, *atypical*, *nonmembers)*. Mean accuracy and mean accurate reaction times were the dependent variables. Reaction times longer than 7000 ms and faster than 200 ms were eliminated from the data analysis (Kiran & Thompson, 2003a). Only correct responses were considered for the reaction time analysis. Item (collapsed across participants) and participant analyses (collapsed across items) were conducted for the results.

RESULTS

Mean accuracy rates

Accuracy rates for the four groups are illustrated in Figure 1. A 3 (category) × 3 (typicality) × 4 (group) ANOVA was performed on the mean accuracy rates as both item (F_1) and subject analysis (F_2). Results revealed a significant main effect for group on the item and subject analysis, $F_1(3, 512)=6.57$, $MS_e=.12$, $p<.0001$, $F_2(3, 297)=2.82$, $MS_e=.07$, $p<.05$, indicating that SI group were significantly more impaired on this task than the young group ($p_1<.0001$, $p_2<.0001$) and the NSI group ($p_1<.0001$, $p_2<.0001$), but not the older group. The three-way interaction between group, category, and typicality was significant only on the item analysis, $F_1(12, 512)=3.01$, $MS_e=.05$, $p<.0001$. Post hoc analysis on the interaction effect revealed that for all four groups, accuracy on atypical *weapons* was significantly worse than all other response types (all effects significant at $p<.001$). In order to identify specific effects of category and typicality within each group, responses were then separated and analysed by group. Bonferroni corrections were applied to all the statistical tests such that only p values lower than .01 were considered statistically significant.

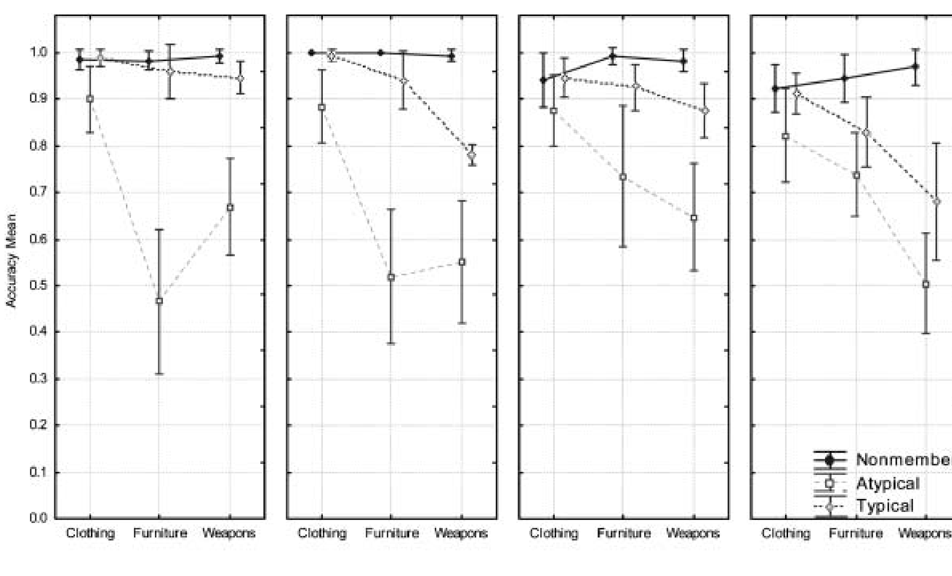

Figure 1. Box and whisker plots showing mean accuracy rates for the four participant groups (young, older, NSI, and SI) across three categories (*clothing*, *furniture*, and *weapons*) for typical, atypical, and nonmember items. Whiskers indicate 95% confidence intervals.

A significant main effect of typicality was observed for all four groups, the young group, $F_1(2, 128)=80.22$, $MS_e=1.36$, $p<.0001$, $F_2(2, 81)=61.69$, $MS_e=.88$, $p<.0001$; older group, $F_1(2, 128)=78.5$, $MS_e=1.46$, $p<.0001$, $F_2(2, 81)=22.9$, $MS_e=.97$, $p<.0001$; NSI group, $F_1(2, 128)=29.38$, $MS_e=.58$, $p<.0001$, $F_2(2, 63)=17.18$, $MS_e=.33$, $p<.0001$; and the SI group, $F_1(2, 128)=32.9$, $MS_e=.73$, $p<.0001$, $F_2(2, 72)=16.03$, $MS_e=.44$, $p<.0001$. For all groups, responses to atypical examples were significantly poorer than typical examples and to nonmembers (all analyses were significant on the item and subject analysis at $p<.001$). Responses to typical and nonmembers were not significantly different.

A significant main effect of category emerged for the young group, $F_1(2, 128)=15.99$, $MS_e=.27$, $p<.0001$, $F_2(2, 81)=12.78$, $MS_e=.18$, $p<.0001$; older group, $F_1(2, 128)=22.53$, $MS_e=.42$, $p<.0001$, $F_2(2, 81)=6.6$, $MS_e=.28$, $p<.001$; NSI group (only on the item analysis), $F_1(2, 128)=4.3$, $MS_e=.08$, $p<.05$; and the SI group, $F(2, 128)=15.07$, $MS_e=.33$, $p<.0001$, $F_2(2, 72)=7.12$, $MS_e=.19$, $p<.01$. Post hoc analyses revealed that for all groups, responses to *clothing* were significantly more accurate than to *weapons*, and for the two normal groups, responses to *clothing* were significantly more accurate than responses to *furniture* as well (all effects were significant on the item and subject analysis at $p<.001$ level).

Finally, interaction effects between typicality and category emerged in all four groups: young group, $F_1(4, 128)=13.19$, $MS_e=.22$, $p<.0001$, $F_2(4, 81)=10.35$, $MS_e=.14$, $p<.0001$; older group, $F_1(4, 128)=10.18$, $MS_e=.18$, $p<.0001$, $F_2(4, 81)=3.09$, $MS_e=.13$, $p<.05$; NSI group, $F_1(4, 128)=3.6$, $MS_e=.07$, $p<.0001$; and SI group, $F_1(4, 128)=6.26$, $MS_e=.18$, $p<.0001$. Post hoc tests revealed that for the two normal groups, atypical *weapons* and atypical *furniture* were significantly poorer than all other response types. Further, older participants responded less accurately to typical *weapons* than typical *clothing*. In contrast, the NSI group responded less accurately to atypical *weapons* compared to all other response types. Finally, for the SI group, responses to atypical *weapons* and typical *weapons* were significantly poorer than all other response types. All effects were significant on the item and subject analysis at $p<.001$.

Mean reaction times

Mean reaction times on the correct responses were calculated by category and typicality for each group (see Figure 2 and Table 2). A 3 (category) × 3 (typicality) × 4 (group) ANOVA was performed on the mean reaction times as item (F_1) and subject analysis (F_2). Results revealed a significant main effect for group, $F_1(3, 512)=163.0$, $MS_e=2.70$, $p<.0001$, $F_2(3, 297)=44.5$, $MS_e=1.63$, $p<.0001$, indicating that reaction times were fastest for the young group followed by the older group, and lastly the two aphasic groups (all effects significant for item and subject analysis at $p<.001$). Reaction times for the two aphasic groups were not significantly different from each other. The three-way group × category × typicality interaction was not significant on the item or subject analysis. Responses were again separated and analysed by group, as the aim of the study was to examine the effect of category and typicality within each participant group. Bonferroni corrections were applied to all the statistical tests such that only p values below .01 were considered statistically significant.

Significant main effects were observed for typicality were observed for the young group, $F_1(2, 128)=24.2$, $MS_e=2127089$, $p<.0001$, $F_2(2, 81)=14.03$, $MS_e=2493679$, $p<.0001$; older group, $F_1(2, 128)=25.5$, $MS_e=3741690$, $p<.0001$, $F_2(2, 81)=10.71$,

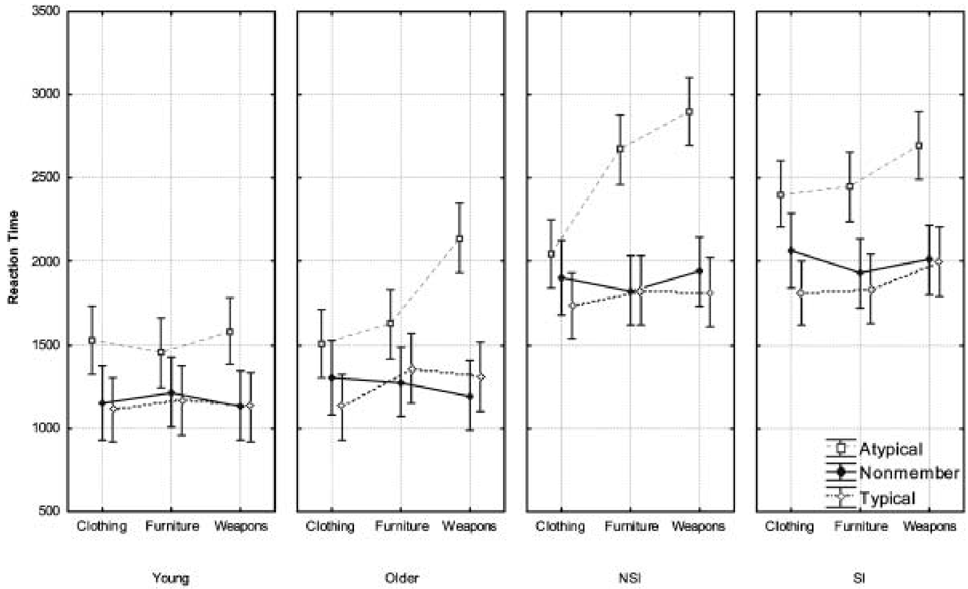

Figure 2. Box and whisker plots showing mean reaction times for the four participant groups (young, older, NSI, and SI) across three categories (*clothing*, *furniture*, and *weapons*) for typical, atypical, and nonmember items. Whiskers indicate 95% confidence intervals.

MS_e=3575336, p<.0001; NSI group, $F_1(2, 128)=39.24$, MS_e=7698890, p<.0001, $F_2(2, 63)=11.49$, MS_e=4258909, p<.0001; and the SI group, $F_1(2, 128)=21.85$, MS_e=5300201, p<.0001, $F_2(2, 72)=4.35$, MS_e=2670450, p<.05. Post hoc analysis for all four groups revealed that response times to atypical examples were significantly longer than typical examples and nonmembers. All effects were significant on the item and subject analysis and at p<.00001 level.

A significant main effect for category was observed for only for the older group, $F_1(2, 128)=4.16$, MS_e=622467, p<.01, and the NSI group, $F_1(2, 128)=6.34$,

TABLE 2
Mean reaction times in milliseconds and standard deviations for category and typicality for the four participant groups

Group	*Category*	*Typical*	*(SD)*	*Atypical*	*(SD)*	*Nonmember*	*(SD)*
	Clothing	1108.2	229.408	1525.13	371.518	1152.14	151.873
Young	Furniture	1168.8	129.879	1454.51	497.072	1215.93	161.633
	Weapons	1128.3	132.545	1582.55	469.614	1131.77	174.646
	Clothing	1125.7	141.97	1507.76	262.43	1304.41	141.068
Older	Furniture	1356.2	228.679	1627.39	394.357	1277.23	309.187
	Weapons	1309.1	258.438	2137.63	915.037	1194.54	182.601
	Clothing	1735.5	254.506	2045.11	285.95	1903.37	333.856
NSI	Furniture	1825.6	281.271	2675.11	728.211	1825.51	214.498
	Weapons	1813.9	252.641	2902.07	812.759	1942.87	330.931
	Clothing	1810.1	341.396	2406.03	538.07	2065.85	381.156
SI	Furniture	1835.4	284.441	2448.84	528.722	1931.39	349.46
	Weapons	1999.5	446.835	2696.49	889.64	2013.44	345.92

MS_e=1244453, p<.01. For both groups, response times to *weapons* were significantly longer than to *clothing* (p_1<.0001). Category effects were not significant for the young group or the SI group. Significant interaction effects between typicality and category were also observed only for the older and the NSI group. For both groups, response times for atypical *weapons* (in addition to atypical *furniture* for the NSI group) were significantly longer than all response types (all effects were significant on the item and subject analysis at p<.001).

Percent typicality effect

To further analyse the difference in reaction times between typical and atypical examples, the advantage for typical examples over the atypical examples was calculated for each group across the three categories. This effect, labelled *percent typicality effect,* was calculated as (Mean Typical – Mean Atypical)/Mean Typical and thus normalised the data across the four groups. A positive value was observed for all four groups, indicating that typical examples were faster than atypical examples across all groups (see Figure 3). The young and older participants demonstrated similar trends, although the typicality effect for *weapons* was larger for the older than the younger group. The NSI group demonstrated a different pattern of the typicality effect across the three categories, the smallest effect was for *clothing* and the largest effect was for *weapons*. Notably, the SI group demonstrated no difference across the three categories in terms of the typicality effect.

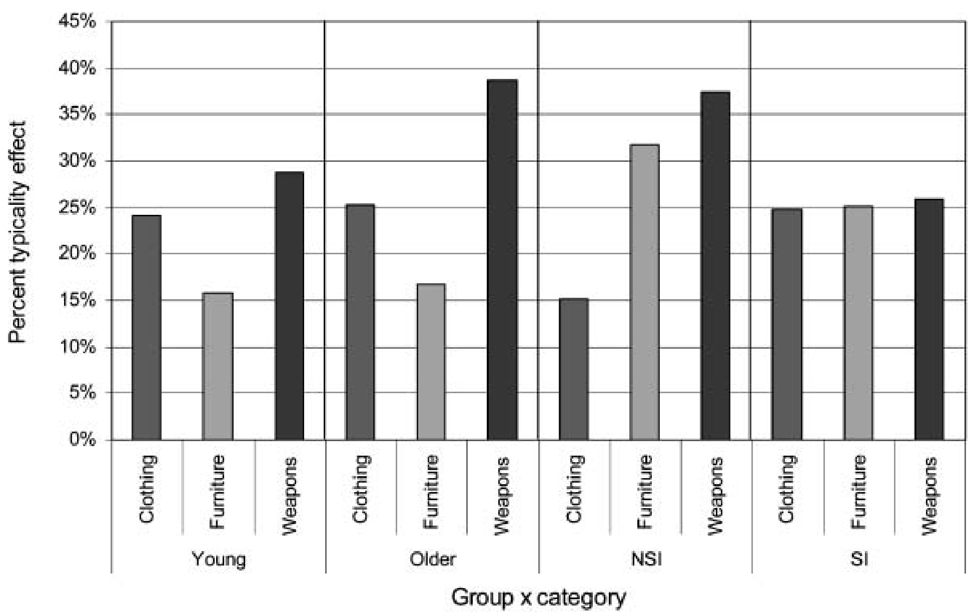

Figure 3. Percentage typicality effect calculated from reaction times using the formula (Mean Typical – Mean Atypical)/Mean Typical for the four participant groups (young, older, NSI, and SI) across three categories (*clothing*, *furniture*, and *weapons*).

TABLE 3
Results of regression analysis for mean accuracy rates and mean response times

Regression results for mean accuracy rates						
	Multiple R^2=0.47		Adjusted R^2=0.30		$F(4, 12)$=2.71	p=.08
Multiple R=0.69	*Beta*	*Std.Err.*	*B*	*Std.Err.*	*t(12)*	*p-level*
Intercept			−0.76	0.61	−1.24	0.24
WAB AQ	0.07	0.36	0.00	0.00	0.19	0.85
WAB AC	−0.13	0.34	−0.12	0.31	−0.39	0.70
PALPA PM Ave **	0.77	0.33	1.82	0.78	2.32	0.04
PALPA SJ ave	−0.09	0.34	−0.09	0.32	−0.28	0.79
Regression results for mean response times						
	Multiple R^2=*0.46*		Adjusted R^2=0.28		$F(4, 12)$=2.56	p=.09
Multiple R=0.68	*Beta*	*Std.Err.*	*B*	*Std.Err.*	*t(12)*	*p-level*
Intercept			−6616.12	5235.85	−1.26	0.23
WAB AQ	−0.62	0.36	−35.31	20.57	−1.72	0.11
WAB AC**	0.81	0.34	6247.44	2627.00	2.38	0.03
PALPA PM Ave	0.61	0.34	12252.41	6681.75	1.83	0.09
PALPA SJ Ave **	−0.89	0.35	−6963.28	2712.26	−2.57	0.02

Regressors are test scores taken from selected WAB and PALPA tests. WAB AQ=WAB Aphasia Quotient, WAB AC=WAB Auditory Comprehension, PALPA PM Ave=PALPA picture–word matching averaged across spoken and written modality, PALPA SJ Ave=PALPA synonym judgement averaged across spoken and written modality. **indicates significant results at $p<.05$

Regression analysis of patient variables on reaction time and accuracy

Since we divided our aphasic participants based on performance on offline semantic measures and not their diagnosis of Broca's or Wernicke's aphasia, we sought to establish the validity of such a distinction criterion. Multiple regression analyses were performed on various patient variables to estimate their predictive power on performance on the online category verification task. These analyses also allowed us to investigate if aphasia severity (as measured by WAB AQ) was a confounding factor in the assignment of participants into NSI and SI groups as discussed in the methods. The regressors entered into the analyses were (a) WAB aphasia quotient, (b) WAB auditory comprehension, (c) PALPA word to picture matching (averaged across spoken and written modalities), and (d) PALPA synonym judgement (averaged across spoken and written modalities). The dependent variables were mean reaction times and mean accuracy.

Results of the regression analyses for mean accuracy and mean reaction times are shown in Table 3. For mean accuracy, the regression was a moderate fit (R^2=47%), but the overall relationship was not significant ($F_{4,12}$=2.71 p=.08). However, PALPA word to picture matching (t_{12}=2.32, p<.05) was a significant predictor of mean response accuracy.[1] Similarly, for mean reaction times, the regression was a moderate fit (R^2=46%), but the overall relationship was not significant ($F_{4,12}$=2.56, p=.09). However, both WAB Auditory comprehension (t_{12}=2.38, p<.05) and

[1] It appears that the small number of participants (N = 17) reduces the statistical significance of the data. As in Table 3, the adjusted r is reduced to 25%. However, we still report the t values as they are significant.

PALPA synonym judgement ($t_{12}=-2.57$, $p<.05$) were significant predictors of participants' reaction times. Not surprisingly, auditory comprehension was a significant predictor of performance. It has been reported in several recent studies that auditory comprehension skills are marked indicators of language recovery and performance in aphasia (Cao, Vikingstad, George, Johnson, & Welch, 1999; Heiss, Kessler, Thiel, Ghaemi, & Karbe, 1999). It should be noted that here aphasia severity (as measured by WAB Aphasia Quotient) was not a significant predictor of performance either on the mean accuracy regression analysis or mean response time regression analysis. These results support our initial claim that performance on specific PALPA subtests (i.e., synonym judgement) was a reliable predictor of reaction time performance on the online category verification task.

DISCUSSION

The aim of this study was to identify differences in activation of typical and atypical examples and nonmembers in three inanimate categories used (*furniture*, *clothing*, and *weapons*) across four experimental groups as measured by accuracy of responses and mean reaction times. Results of this experiment demonstrated that in general, typical examples were processed faster and more accurately than atypical examples within each category across the four participant groups. Of interest is that among the three categories presented, *clothing* was processed faster and more accurately than *weapons*. A category-specific effect within inanimate categories was not hypothesised and its presence in the context of the typicality effect has implications for the way in which different categories are processed in the brain. This study was different in that participants were assigned to two experimental groups based on their performance on offline tests of semantic processing rather than individual diagnosis of aphasia type. Results from a regression analysis tentatively supported this classification, as specific offline semantic processing tests such as synonym judgement appeared to be a reliable predictor of reaction time performance on the online verification task.

Analysis of accuracy rates on typicality and category across the four groups revealed the following results. First, the SI group performed worse on this task than any other participant group, indicating that, as predicted, the impairments the SI participants demonstrated on the synonym judgement tasks were associated with impaired performance on the category verification task. Second, all groups made more errors on atypical than typical items and nonmembers. The older and SI groups were also less accurate on typical examples than nonmembers. The magnitude of difference in accuracy between nonmembers and typical examples was much smaller for the older controls (99% accuracy for nonmembers and 90% accuracy for typical examples) than for the SI group (94% accuracy for nonmembers and 81% for typical examples). This difference in accuracy rates of typical versus nonmembers has not been found in other studies where the participants were divided according to lesion site (anterior vs posterior) (Grober et al., 1980) or type of aphasia (Broca's vs Wernicke's) (Kiran & Thompson, 2003a), and is possibly reflective of the nature of the stimuli than of the experimental groups studied.

Third, the accuracy analysis also revealed that all the participants made more errors with *weapons* than *furniture* and *clothing*. However, when the data from the four groups were analysed separately, it became apparent that for both the normal groups, accuracy for atypical *clothing* was significantly better than atypical *furniture* or *weapons*, whereas for the two aphasic groups accuracy for typical and atypical *weapons* were significantly impaired. Only inanimate categories and their examples were selected in this experiment since, as discussed in the introduction, differential processing between animate and inanimate categories has been observed in several studies. However, it seems that within inanimate categories some categories are processed differently from others. One seemingly simple explanation for this finding may be the inherent familiarity of the category, i.e., examples of *weapons* are less familiar than examples of *furniture* or *clothing*. However, the stimuli in the present experiment were controlled for familiarity across the three categories. Nevertheless, a similar finding was reported by Kiran and Thompson (2003a) in the domain of animate categories, where all their experimental groups (normal young, normal older, Broca's, Wernicke's) were less accurate with examples of the category *fish* than items of the two other categories (*birds*, *vegetables*). As suggested by Malt and Smith (1982), it may be possible that the overall familiarity of the category, not reflected in the individual item familiarity, may have influenced the category membership decisions in the present study.

Evidence from mean reaction time data, in general, supported the accuracy data. As expected, participants with aphasia were slower on the task than their normal young and age-matched controls. However, across the four groups, typicality effects were more robust than category effects on the reaction time analysis. First, all four groups demonstrated slower reaction times for atypical examples compared to typical examples and nonmembers. Differences between typical and nonmembers were not significant. Second, only the older group and the NSI group demonstrated significant category effects, i.e., faster reaction times for *clothing* compared to *weapons*. Further analysis in these two participant groups revealed that reaction times for atypical *weapons* examples were significantly longer than any other example type, a finding resonant with the accuracy data.

At first pass, it appears as though the SI group had a similar performance to the young control group in terms of reaction time data. That is, both the young and the SI group demonstrated only significant typicality effects and no category effects. However, calculation of the percentage typicality effect normalised across the four groups revealed different results (see Figure 3). The percentage typicality effect of the young controls was similar to their older counterparts across the three categories. Specifically, the percentage typicality effect was the least for *furniture* and the most for *weapons* in both groups. Contrastingly, there was no preferential advantage for typical examples in any of the categories in the SI group, a finding different from the remaining three participant groups.

One potential explanation for this finding may be that these participants made so many errors on the typical and atypical examples within each category that subsequent exclusion of incorrect responses eliminated any category-specific advantage for typical members of *weapons* and *furniture*. In order to explore this possibility, partial correlations (controlling for category) for each participant group were performed on the mean accuracy rates and the average typicality rating expressed in average z scores. For the controls, a high negative correlation between typicality and error proportion would be predicted, indicating that the lower the average z score of the example (the more typical the example) the higher its accuracy. For the SI group, however, the strength of this correlation should be weakened or absent, since participants were presumably making numerous errors even on the typical examples. Figure 4 illustrates the results of this analysis. Significant correlations were observed for all four groups: young group ($r^2 = -.66$, $p < .05$), older

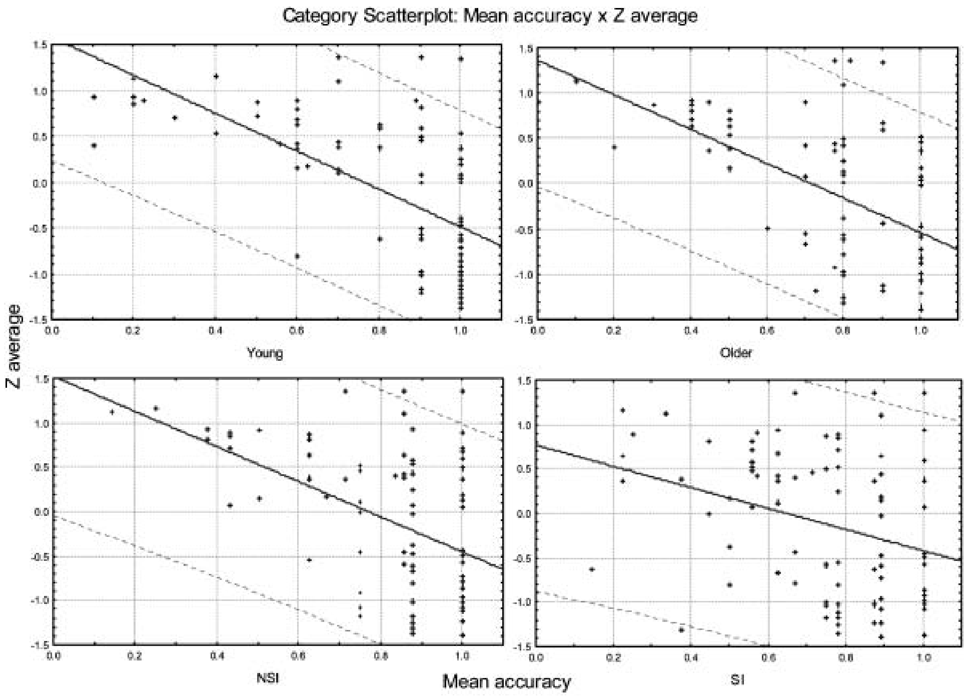

Figure 4. Category scatterplots showing the spread of data across the four participant groups (young, older, NSI, and SI). Z average ratings range from −1.5 (very typical) to +1.5 (very atypical), whereas mean accuracy ranges from 0% to 100%. Data were derived from partial correlations (controlling for category) between mean accuracy rates and *z* average scores. Vertical lines indicate linear fits of the data and dashed lines indicate 95% prediction ranges.

($r^2 = -.69, p < .05$), NSI group ($r^2 = -.51, p < .05$), and SI group ($r^2 = -.40, p < .05$).[2] As predicted, the strength of the correlation was the least for the SI group and the most for the young and older control group. These results suggested that SI participants were indeed impaired in their representation of typical examples of the category, eliminating the category-specific advantage demonstrated by other participant groups. Given that the SI group demonstrated a robust typicality effect despite an ostensible impairment in representation of typical examples, it can be posited that typicality effects can arise in a disrupted network, weakening associations at all levels of typicality but not disrupting the relative strengths of exemplars within a concept. A similar hypothesis has been put forth by Johnson and colleagues (Johnson, Herman, & Bonilla, 1995) in the representation of category members in participants with Alzheimer's disease.

Comparison of the performance between NSI and SI groups revealed some interesting observations. First, the SI group made significantly more errors on atypical and typical examples than their NSI counterparts. These results suggest that the deficits in the SI group may reflect impairment in the representation of semantic categories since even typical examples are affected. Contrary to our predictions, reaction time patterns were similar across both groups, suggesting that perhaps the

[2] It should be noted, as can be seen in Figure 4, that accuracy rates in general were at ceiling levels, but a significant correlation still emerged.

overall speed of lexical processing was similarly affected following brain damage. It is important to note that both groups constituted a spectrum of aphasia type including anomic, Broca's, conduction, and transcortical motor aphasia, perhaps contributing to the lack of quantitative differences between the two groups. Despite the similarity in reaction times, it is concluded that the above-mentioned differences between the two groups were reflective of their performance on the online and offline semantic processing tasks.

The results of the present study are mostly consistent with findings of a previous experiment (Kiran & Thompson, 2003a). Both experiments demonstrated the robustness of the typicality effect across participant groups and across semantic categories, and differences between individuals with brain damage and their normal controls. Whereas the previous study examined differences between participants with Broca's and Wernicke's aphasia and found Wernicke's aphasic patients to be impaired in the representation of typicality, the present study found the SI group to be worse on the task than the NSI group although both groups showed robust typicality effects.

Some additional distinctions emerged in the nature and magnitude of typicality effects between the two studies that are worth mentioning. First, reaction times in the present study were generally slower than the previous study that examined animate categories. For instance, the fastest reaction times in the present study (young typical = 1108 ms) were generally longer than those in the Kiran and Thompson study (young typical = 869 ms). Related to the previous point, the magnitude of the typicality effect, which illustrated the advantage of typical examples over atypical examples normalised across groups, was also generally higher in the present study across all participant groups.

Taken together, the above findings point to a general difference in the representation and access of animate and inanimate categories, supporting the previous discussion that animate categories involve more absolute category judgements than inanimate categories, which entail relatively more graded processing (Barton & Komatsu, 1989; Diesendruck & Gelman, 1999; Estes, 2003). Correspondingly, typicality effects (i.e., graded category representation) were larger for inanimate categories than for animate categories in our experiments. It is possible that category boundaries for inanimate categories are fuzzier than animate categories, partly accounting for their gradedness.

CONCLUSION

To conclude, the present experiment demonstrated the typicality effect in normal individuals and in individuals with aphasia. These results are consistent with recent work in our lab showing typicality effects during online feature verification tasks (Kiran & Allison, 2006). Further, differences emerged in the processing of categories, where in responses to *clothing* were more accurate than responses to *furniture* or *weapons*. These results contribute to the continuing debate on category specificity and the influence of brain damage on this organisation (Devlin et al., 2002; Laws & Neve, 1999; Tyler, Durrant-Peatfield, Levy, Voice, & Moss, 1996). These findings also have implications for rehabilitation of aphasia, where naming treatments that are based on semantic feature analysis (Kiran & Thompson, 2003b) should take into account semantic features accessed for the predominantly form-based animate categories (e.g., *has legs*) compared to the predominantly function-based inanimate categories (Devlin et al., 2002).

Manuscript received 10 November 2005
Manuscript accepted 7 April 2006
First published online 21 June 2007

REFERENCES

Ashcraft, M. H. (1978). Property norms for typical and atypical items from 17 categories: A description and discussion. *Memory & Cognition*, *6*(3), 227–232.

Barr, R. A., & Caplan, L. J. (1987). Category representations and their implications for category structure. *Memory & Cognition*, *15*(5), 397–418.

Barton, M. E., & Komatsu, L. K. (1989). Defining features of natural kinds and artifacts. *Journal of Psycholinguistic Research*, *18*(5), 433–447.

Basso, A., Lecours, A. R., Moraschini, S., & Vanier, M. (1985). Anatomoclinical correlations of the aphasias as defined through computerized tomography: Exceptions. *Brain and Language*, *26*(2), 201–229.

Bruner, J. S., Goodnow, J., & Austin, G. (1956). *A study of thinking*. New York: Wiley.

Cabeza, R., & Nyberg, L. (2000). Imaging cognition II: An empirical review of 275 PET and fMRI studies. *Journal of Cognitive. Neuroscience*, *12*(1), 1–47.

Cao, Y., Vikingstad, E. M., George, K. P., Johnson, A. F., & Welch, K. M. (1999). Cortical language activation in stroke patients recovering from aphasia with functional MRI. *Stroke*, *30*(11), 2331–2340.

Casey, P. J. (1992). A reexamination of the roles of typicality and category dominance in verifying category membership. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *18*(4), 823–834.

Cole-Virtue, J., & Nickels, L. (2004). Spoken word to picture matching from PALPA: A critique and some new matched sets. *Aphasiology*, *18*(2), 77–102.

Collins, A. M., & Loftus, E. F. (1975). A spreading-activation theory of semantic processing. *Psychological Review*, *82*(6), 407–428.

Del Toro, J. F. (2000). An examination of automatic versus strategic semantic priming effects in Broca's aphasia. *Aphasiology*, *14*(9), 925–947.

Devlin, J. T., Russell, R. P., Davis, M. H., Price, C. J., Moss, H. E., & Fadili, M. J., et al. (2002). Is there an anatomical basis for category-specificity? Semantic memory studies in PET and fMRI. *Neuropsychologia*, *40*(1), 54–75.

Diesendruck, G., & Gelman, S. A. (1999). Domain differences in absolute judgements of category membership: Evidence for an essentialist account of categorisation. *Psychonomic Bulletin & Review*, *6*(2), 338–346.

Estes, Z. (2003). Domain differences in the structure of artifactual and natural categories. *Memory & Cognition*, *31*(2), 199–214.

Frances, N., & Kucera, H. (1982). *Frequency analysis of English usage*. Boston, MA: Houghton Mifflin.

Garrard, P., Lambon Ralph, M. A., Hodges, J. R., & Patterson, K. (2001). Prototypicality, distinctiveness, and intercorrelation: Analyses of the semantic attributes of living and nonliving concepts. *Cognitive Neuropsychology*, *18*(2), 125–174.

Gennari, S., & Poeppel, D. (2003). Processing correlates of lexical semantic complexity. *Cognition*, *89*(1), B27–41.

Grober, E., Perecman, E., Kellar, L., & Brown, J. (1980). Lexical knowledge in anterior and posterior aphasics. *Brain and Language*, *10*(2), 318–330.

Grossman, M. (1981). A bird is a bird is a bird: Making reference within and without superordinate categories. *Brain and Language*, *12*(2), 313–331.

Hampton, J. (1993). Prototype models of concept representation. In I. Van Mechelen, J. A. Hampton, & P. T. R. S. Michlanski (Eds), *Categories and concepts: Theoretical views and inductive data analysis*. London: Academic Press.

Hampton, J. A. (1979). Polymorphous concepts in semantic memory. *Journal of Verbal Learning & Verbal Behavior*, *18*(4), 441–461.

Hampton, J. A. (1995). Testing the prototype theory of concepts. *Journal of Memory & Language*, *34*(5), 686–708.

Heiss, W. D., Kessler, J., Thiel, A., Ghaemi, M., & Karbe, H. (1999). Differential capacity of left and right hemispheric areas for compensation of poststroke aphasia. *Annals of Neurology*, *45*(4), 430–438.

Hough, M. S. (1993). Categorisation in aphasia: Access and organisation of goal-derived and common categories. *Aphasiology*, *7*(4), 335–357.

Johnson, M. K., Herman, A. H., & Bonilla, J. L. (1995). Semantic relations and Alzheimer's disease: Typicality and direction of testing. *Neuropsychology*, *9*(4), 529–536.

Kaplan, E., Goodglass, H., & Weintraub, S. (2001). *Boston Naming Test* (2nd ed.). Philadelphia: Lippincott Williams & Wilkins.

Kay, J., Lesser, R., & Coltheart, M. (1992). *The Psycholinguistic Assessment of Language Processing in Aphasia (PALPA)*. Hove, UK: Lawrence Erlbaum Associates Ltd.

Keil, F. C. (1989). *Concepts, kinds, and cognitive development*. Cambridge, MA: The MIT Press.

Kertesz, A. (1982). *The Western Aphasia Battery*. Philadelphia: Grune & Stratton.

Kiran, S., & Allison, K. (2006). Effect of typicality on lexical processing of feature based and perceptually based categories. *Manuscript submitted for publication*.

Kiran, S., & Thompson, C. K. (2003a). Effect of typicality on online category verification of animate category exemplars in aphasia. *Brain and Language*, *85*(3), 441–450.

Kiran, S., & Thompson, C. K. (2003b). The role of semantic complexity in treatment of naming deficits: Training semantic categories in fluent aphasia by controlling exemplar typicality. *Journal of Speech, Language and Hearing Research*, *46*(4), 773–787.

Kiran, S. (in press). Semantic complexity in the treatment of naming deficits. *American Journal of Speech Language Pathology*.

Kiran, S. D. (2002). Effect of exemplar typicality on naming deficits in fluent aphasia. *Dissertation Abstracts International: Section B: The Sciences & Engineering*, *62*(11-B), 5076. US: Univ. Microfilms International.

Larochelle, S., & Pineau, H. (1994). Determinants of response times in the semantic verification task. *Journal of Memory & Language*, *33*(6), 796–823.

Laws, K., & Neve, C. (1999). A normal category specific advantage for naming living things. *Neuropsychologia*, *37*, 1263–1269.

Malt, B. C., & Smith, E. E. (1982). The role of familiarity in determining typicality. *Memory & Cognition*, *10*(1), 69–75.

McCloskey, M. E. (1980). The stimulus familiarity problem in semantic memory research. *Journal of Verbal Learning and Verbal Behavior*, *19*, 485–502.

McCloskey, M. E., & Glucksberg, S. (1978). Natural categories: Well defined or fuzzy sets? *Memory & Cognition*, *6*(4), 462–472.

McRae, K., de Sa, V. R., & Seidenberg, M. S. (1997). On the nature and scope of featural representations in word meaning. *Journal of Experimental Psychology General*, *126*, 99–130.

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

Milberg, W., Blumstein, S. E., & Dworetzky, B. (1987). Processing of lexical ambiguities in aphasia. *Brain and Language*, *31*(1), 138–150.

Posner, M. I., & Keele, S. W. (1968). On the genesis of abstract ideas. *Journal of Experimental Psychology*, *77*, 353–363.

Rips, L. J., Shoben, E. J., & Smith, E. E. (1973). Semantic distance and the verification of semantic distance. *Journal of Verbal Learning and Verbal Behavior*, *12*, 1–20.

Rosch, E. (1973). On the internal structure of perceptual and semantic categories. In T. E. Moore (Ed.), *Cognitive development and the acquisition of language*. New York: Academic Press.

Rosch, E. (1975). Cognitive representations of semantic categories. *Journal of Experimental Psychology: General*, *104*(3), 192–233.

Rosch, E., & Mervis, C. B. (1975). Family resemblances: Studies in the internal structure of categories. *Cognitive Psychology*, *7*(4), 573–605.

Smith, E. E., Shoben, E. J., & Rips, L. J. (1975). Structure and process in semantic memory: A featural model of semantic association. *Psychological Review*, *81*, 214–241.

Thompson-Schill, S. L. (2003). Neuroimaging studies of semantic memory: Inferring "how" from "where". *Neuropsychologia*, *41*(3), 280–292.

Tyler, L. K., Durrant-Peatfield, M. R., Levy, J. P., Voice, J. K., & Moss, H. E. (1996). Distinctiveness and correlations in the structure of categories – behavioural data and connectionist model. *Brain and Language*, *55*, 89–92.

Tyler, L. K., & Moss, H. E. (2001). Towards a distributed account of conceptual knowledge. *Trends in Cognitive Science*, *5*(6), 244–252.

Tyler, L. K., Moss, H. E., Durrant-Peatfield, M. R., & Levy, J. P. (2000). Conceptual structure and the structure of concepts: A distributed account of category-specific deficits. *Brain and Language*, *75*(2), 195–231.

Uyeda, K. M., & Mandler, F. (1980). Prototypicality norms for 28 semantic categories. *Behavior Research Methods and Instrumentation*, *12*(6), 587–595.

Vanoverberghe, V., & Storms, G. (2003). Feature importance in feature generation and typicality rating. *European Journal of Cognitive Psychology*, *15*(1), 1–18.

Wilson, D. (1988). *MRC Psycholinguistic Database. Machine Usable Dictionary*. Retrieved, from http://www.psy.uwa.edu.au/mrcdatabase/uwa_mrc.htm).

APPENDIX

TYPICAL AND ATYPICAL EXAMPLES AND NONMEMBERS USED FOR THE THREE CATEGORIES IN THE SEMANTIC PRIMING TASK

Clothing			*Furniture*			*Weapons*		
Typical	*Atypical*	*Nonmembers*	*Typical*	*Atypical*	*Nonmembers*	*Typical*	*Atypical*	*Nonmembers*
Blouse	Apron	Apricot	Bed	Carpet	Apple	Axe	Brick	Broccoli
Boots	Bathing suit	Banana	Bookcase	Chandelier	Asparagus	Bayonet	Candlestick	B uffalo
Jacket	Belt	Cabbage	Bureau	Drapes	Bear	Bazooka	Dynamite	Camel
Jeans	Bib	Cow	Cabinet	Furnace	Beet	Bomb	Ice pick	Cantaloupe
Overalls	Cape	Elephant	Chair	Hammock	Blueberry	Dagger	Laser	Carrots
Pajamas	Earmuffs	Kangaroo	Coffee table	Mirror	Cauliflower	Grenade	Poison	Cat
Pants	Flight suit	Onion	Desk	Picture	Corn	Knife	Razor	Celery
Shirt	Garter	Peach	Dresser	Pillow	Frog	Machine gun	Rock	Grape
Shorts	Gloves	Pepper	End table	Radio	Horse	Missile	Rocket	Lettuce
Skirt	Helmet	Pineapple	Footstool	Rug	Mango	Revolver	Rope	Pig
Suit	Hood	Spinach	Lamp	Swing	Monkey	Rifle	Scissors	Rabbit
Sweater	Rainwear	Cherry	Loveseat	Toy box	Mushroom	Shotgun	Slingshot	Raspberry
Sweat suit	Slippers	Turtle	Nightstand	Trunk	Orange	Spear	Stick	Strawberry
Underwear	Suspender	Tomato	Recliner	Umbrella stand	Pears	Sword	Submarines	Watermelon
Vest	Thong	Zebra	Sofa	Wastebasket	Snake	Torpedo	Tanks	Radish