Is it a letter? Is it a number? Processing of numbers within SMS shortcuts

LESYA Y. GANUSHCHAK, ANDREA KROTT, AND ANTJE S. MEYER
University of Birmingham, Birmingham, England

For efficiency reasons, words in electronic messages are sometimes formed by combining letters with numbers, as in *gr8* for "great." The aim of this study was to investigate whether a digit incorporated into a letter–digit shortcut would retain its numerosity. A priming paradigm was used with letter–digit shortcuts (e.g., *gr8*) and matched pseudoshortcuts (e.g., *qr8*) as primes. The primes were presented simultaneously with sets of dots (targets) for which even/odd decisions were required, or they appeared 250 msec before target onset. When pseudoshortcuts were presented, decision latencies were shorter when the target and the digit in the prime were matched in parity than when they were mismatched. This main effect of match was not significant for shortcuts. The results suggest that the number concepts of digits combined with letters become activated but are quickly suppressed or deactivated when the digit is part of an existing shortcut.

At the time of *Sesame Street*, everything was easy: *l* was a letter, and *8* was a number. However, as the popularity of sending messages via mobile phones and e-mail increased, so did the necessity of conveying messages more efficiently. For efficiency reasons, words in electronic messages are sometimes written as shortcuts by combining letters with numbers, as in *gr8*, *4ever*, or *l8r*. The increased popularity of abbreviations and acronyms in daily communication demands more research into their mental representation and their effects on language use and diachronic language development. Recently, Perea, Acha, and Carreiras (2009) investigated how readers processed sentences written in short message service (SMS) language. They recorded the eye movements of participants reading sentences written in SMS language (e.g., *my hols wr gr8*) and conventionally written sentences (e.g., *my holidays were great*). Perea and colleagues showed that reading shortcuts was more effortful than reading conventionally written words, as indicated by longer reading times and more fixations. This suggests that processing shortcuts can be a slow and demanding task for a reader.

Some common shortcuts (e.g., *gr8*) include digits. One reason why these shortcuts can take longer to process is that the number concept associated with the digit might hinder the processing of the words. The goal of the present study was to investigate whether number concepts are automatically activated when digits are combined with letters to form words, as in *gr8*. Some of these shortcuts (e.g., *2day*) resemble compound words, which are combinations of words (e.g., *bedroom < bed + room*), whereas others (e.g., *h8*) resemble words that include semantically unrelated embedded words (e.g., *hatch* with embedded *hat*). It has been shown that, during the processing of a compound word, the constituent morphemes initially be-

come semantically activated, even if they are irrelevant for the understanding of the whole word (e.g., *hogwash < hog + wash*; Libben, Gibson, Yoon, & Sandra, 2003; Zwitserlood, 1994), and are then suppressed or inhibited by the activation of the stored semantic representation of the whole word (Libben, 1998). A loss of such inhibitory processes can be seen in some aphasic patients who simultaneously activate conceptual representations of constituents and whole words (e.g., *yellowbelly < a yellow stomach . . . a chicken*; Libben, 1998). There is also evidence that, during the visual processing of monomorphemic words, the meanings of embedded words (e.g., *hat* in *hatch*) become activated (e.g., Bowers, Davis, & Hanley, 2005; Rodd, 2004).

Embedded digits have also been shown to activate their meanings. For example, a study using the numerical Stroop paradigm, in which participants compared simultaneously presented Arabic digits with respect either to their printed size or to the indicated magnitude, showed automatic access to magnitude information both in adults (e.g., Pinel, Piazza, Le Bihan, & Dehaene, 2004) and in children (e.g., Rubinstein, Henik, Berger, & Shahar-Shalev, 2002). Also, judging *47*, for example, as being smaller than *62* has been shown to take longer than judging *52* as being smaller than *67* because, in the former case, the decision is hindered by the incongruence between the rightmost constituents (*7* and *2*) and the whole numbers (*47* and *62*; Nuerk, Weger, & Willmes, 2001; but see Ganor-Stern, Pinhas, & Tzelgov, 2009; Zhou, Chen, Chen, & Dong, 2008).

Thus, in both the psycholinguistic and the number processing literatures, there is evidence that suggests that, during the processing of shortcuts with embedded digits, the meanings of the digits might get activated. However, there are also some indications that embedded digits

---

L. Y. Ganushchak, a.y.ganushchak@bham.ac.uk

might simply be encoded in a letter-like manner. For instance, Perea, Duñabeitia, and Carreiras (2008) showed that words preceded by primes that included digits (e.g., M4T3R14L–MATERIAL) were responded to as fast as words preceded by identity primes (e.g., MATERIAL–MATERIAL) were (see also Carreiras, Duñabeitia, & Perea, 2007; Dehaene & Cohen, 2007; Perea et al., 2009). In those studies, however, digits replaced visually similar letters. In the present study, we considered the processing of shortcuts in which digits replaced visually dissimilar but homophonous parts of words or phrases, as in *gr8*.

We can see two ways in which shortcuts with embedded numbers might be processed. First, it may be that readers assemble a shortcut's phonological form by retrieving and combining the phonological forms of the digits and the letters. Second, skilled readers of shortcuts might additionally have stored orthographic representations for entire shortcuts (similar to stored representations of orthographically irregular words) (see Frost, 1998; Rastle & Brysbaert, 2006), which would be linked directly to the associated sound forms and meanings. This would constitute two processing routes similar to those in Coltheart and colleagues' (Coltheart, Rastle, Perry, Langdon, & Ziegler, 2001) reading model, which also distinguishes a grapheme–phoneme conversion route and a full-form access route. There is recent evidence that everyday acronyms like STATS, FBI, and WC have their own entries in the mental lexicon (Brysbaert, Speybroeck, & Vanderelst, 2009). It is therefore plausible that SMS shortcuts have their own lexical representations as well. But even if a shortcut is accessed through a full-form route, the number concept might still be activated via the grapheme-to-phoneme conversion route.

Regardless of which of these two routes readers take, the digits in shortcuts could function exclusively in a letter-like fashion (i.e., being linked to a phonological form but not to a semantic representation). But if they have a dual character, they could function in a letter-like fashion, while also activating the associated number concept. The goal of the present study was to discriminate between these two options—that is, to determine whether or not digits in shortcuts activate number concepts.

We used a task that involved parity (even/odd) judgments on sets of up to four dots (the targets). The decisions had to be made in the presence of primes, which were letter–digit shortcuts (e.g., *gr8*) and letter–digit pseudoshortcuts (e.g., *qr8*). In addition, we manipulated the stimulus onset asynchrony (SOA) between prime and target and the parity match or mismatch between the digit embedded in the prime and that embedded in the target. We presented the prime simultaneously with the target (with an SOA of 0 msec) or starting 250 msec before target onset (with an SOA of −250 msec). For target sets of up to four objects, the numerosity can be apprehended immediately—without the need to count—through a process called *subitizing* (Logan & Zbrodoff, 2003; Revkin, Piazza, Izard, Cohen, & Dehaene, 2008). Access to lexical information, on the other hand, might take a few hundred milliseconds (Barber & Kutas, 2007).

It is possible that the number concept of a digit in a shortcut will be briefly activated but suppressed a little later when the meaning of the shortcut is activated. For pseudoshortcuts, we expected that, at both SOAs, responses would be faster when the embedded digits matched the parity of the dots than when they did not (e.g., Fabre & Lemaire, 2005; Reynvoet, Caessens, & Brysbaert, 2002). The question was whether this effect would also occur for shortcuts. If digits activate number concepts in the same way in shortcuts as in pseudoshortcuts, this should be the case. If number concepts are not activated at all when shortcuts are processed, we would not expect the parity of the digits to affect the responses to the targets. Finally, if the number concepts are initially activated but are deactivated when the meaning of the whole shortcut becomes activated, a parity-match effect might be present at SOA 0, but not at SOA −250 msec.

METHOD

Participants

Forty students of the University of Birmingham (37 female) took part in the experiment (mean age = 19.3 years, $SD = 0.8$). Participants were assigned to two groups: SOA 0 and SOA −250. Due to technical problems, the data from 1 participant were excluded from the analysis. All participants were right-handed native speakers of English and had normal or corrected-to-normal vision. Participants gave written informed consent prior to participating in the study. They received course credits for their participation.

Materials and Design

The targets were single dots or sets of two, three, or four dots. They were presented in a single row in the center of the screen. Each target was used on 44 trials and was accompanied by a prime, which was either one of 22 letter–digit shortcuts or one of 22 matched letter–digit pseudoshortcuts (see the Appendix). Pseudoshortcuts were created from shortcuts by replacing one letter (e.g., *gr8*–*qr8*). Note that digits were never replaced. The primes appeared slightly above, below, or to the side of the targets (see Figure 1). In order to make it difficult for the participants to ignore the primes, the primes' position varied randomly from trial to trial. Each prime appeared four times, once in combination with each target.

The experiment consisted of parity match and mismatch trials. On match trials, the number in the shortcut and the number of dots were either both even or both odd. On mismatch trials, the number in the shortcut was even and the number of dots was odd or vice versa (see Figure 1). For example, *gr8* in combination with one or three dots formed mismatch trials, and *gr8* in combination with two or four
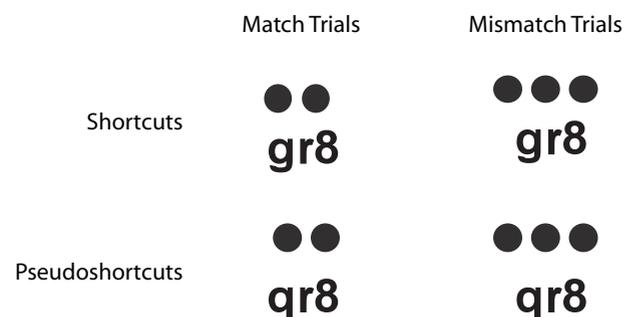
Figure 1. Schematic representation of the displays shown on match and mismatch trials.

dots formed match trials. In total, there were 88 match trials and 88 mismatch trials. In addition to the experimental trials, there were 12 practice trials featuring the four targets, each presented three times together with two shortcuts and two pseudoshortcuts that did not occur in experimental trials.

Procedure

The experiment was controlled by the software package E-Prime 1.2 (Schneider, Eschman, & Zuccolotto, 2002). Stimuli were presented using a Samsung SyncMaster 753s monitor. Participants' responses were recorded using a Cedrus RB-530 response pad.

Each trial started with the presentation of a fixation point in the center of the screen for a duration varying between 500 and 1,000 msec (M = 750 msec). At SOA 0, the fixation cross was followed by the prime and target, which appeared simultaneously and remained in view until the participant responded. At SOA -250, primes appeared 250 msec prior to target onsets. In both cases, primes and targets remained on the screen until the participant responded. After 500 msec, the next trial began. The hand for parity responses (left or right hand for even responses) was counterbalanced across participants. After the experiment, participants were given a listing of all shortcuts used in the experiment and were asked to indicate for each shortcut how often (never, once a year, once a month, once a week, every 2 days, once a day, or several times a day) they read and wrote it, and to write down its meaning (adapted from Balota, Pilotti, & Cortese, 2001).

RESULTS

The participants' modal response in the questionnaire was that they read and wrote the shortcuts once a week. For each participant, any shortcuts they did not know the meaning of were removed from the analysis. This was the case for, on average, 1.4 shortcuts (SD = 1.17) per participant. The error rate for the parity judgments was very low (0.55%), and no error analysis was conducted.

Latency analyses were performed on correct trials only. Latencies shorter than 300 msec or longer than 1,500 msec were excluded from the analyses. The condition means are displayed in Table 1. Separate analyses were carried out on participant and item means, yielding F_1 and F_2 statistics, respectively, and did not show significant interactions involving SOA [for match × SOA, both Fs < 1; for prime type × SOA, F_1 < 1, and $F_2(1,41)$ = 3.04, p = .19; for match × prime type × SOA, both Fs < 1]. There was a significant main effect of match, with participants responding significantly faster (by 20 msec) in the match than in the mismatch condition [$F_1(1,37)$ = 9.60, p = .004; $F_2(1,41)$ = 9.09, p = .004]. There was also a main effect of SOA [$F_1(1,37)$ = 10.06, p = .003; $F_2(1,41)$ =

159.94, p = .001], with participants responding faster (by 56 msec) at SOA -250 than at SOA 0. There was no main effect of prime type (both Fs < 1). However, there was a significant interaction of prime type and match [$F_1(1,37)$ = 4.69, p = .03; $F_2(1,41)$ = 7.62, p = .009].

Further analysis showed a main effect of match for pseudoshortcuts [$F_1(1,37)$ = 13.39, p = .001; $F_2(1,41)$ = 15.79, p = .001], with latencies being shorter by 30 msec in the match than in the mismatch condition, but not for shortcuts [$F_1(1,37)$ = 1.32, p = .26; F_2 < 1], for which latencies in the match and mismatch conditions differed by only 9 msec. In other words, when the primes were pseudoshortcuts, participants were faster on match trials than on mismatch trials; but when they were existing shortcuts, the embedded digits did not significantly affect the parity judgments on the targets.

No interactions involving SOA were significant [for match × SOA, both Fs < 1; for prime type × SOA, F_1 < 1 and $F_2(1,41)$ = 3.04, p = .19; for match × prime type × SOA, both Fs < 1]. However, the pattern of response times suggests that there was an effect of match for both shortcuts and pseudoshortcuts at SOA 0, but only for pseudoshortcuts at SOA -250.

Separate analyses for each SOA confirm this impression: At SOA 0, there was a main effect of match only [$F_1(1,19)$ = 6.11, p = .02; $F_2(1,21)$ = 18.57, p < .001]. There was no interaction of prime type and match (both Fs < 1) and no main effect of prime type [F_1 < 1; $F_2(1,21)$ = 2.64, p = .12]. Thus, at this SOA, digits embedded in shortcuts and in pseudoshortcuts affected the processing of the targets in the same way. By contrast, at SOA -250, there was a significant interaction of prime type and match [$F_1(1,18)$ = 4.79, p = .04; $F_2(1,21)$ = 7.61, p = .01], but no main effect of match [$F_1(1,18)$ = 3.76, p = .07; $F_2(1,21)$ = 2.10, p = .16] or prime type (both Fs < 1). Analyses of simple effects yielded a main effect of match for pseudoshortcuts [$F_1(1,18)$ = 8.66, p = .01; $F_2(1,21)$ = 5.96, p = .02], with latencies being shorter by 34 msec in the match than in the mismatch condition, but not for shortcuts (both Fs < 1).

DISCUSSION

The goal of the present study was to investigate whether digits embedded in pseudoshortcuts and shortcuts were processed as numbers or functioned in the same way as letters—that is, without activating any conceptual information. We found a significant main effect of match and a significant interaction of prime type and match. Follow-up analyses showed that there was a significant match effect for pseudoshortcuts, but not for shortcuts. These results suggest that the digits activated the associated number concepts only and affected the categorization of the sets of dots when they appeared in pseudoshortcuts, but not when they appeared in shortcuts.

Our results stand in contrast to the literature that has shown that words and numbers embedded in other words are activated at least briefly (e.g., Bowers et al., 2005; Libben et al., 2003; Nuerk et al., 2001; Rodd, 2004; Zwit-

Table 1
Mean Reaction Times and Standard Deviations
(in Milliseconds) As a Function of Match and Prime Type

Participant Group	Condition	Prime Type			
		Shortcuts		Pseudo-shortcuts	
		M	SD	M	SD
SOA 0	Match	613	69	603	79
	Mismatch	627	61	630	74
SOA -250	Match	559	49	547	52
	Mismatch	563	63	581	47

serlood, 1994). Given the previous findings, it is surprising that the number concepts associated with the digits in shortcuts appear not to have been activated at all. However, at SOA 0, there was a 14-msec difference between the matching and mismatching shortcut conditions, and at this SOA, the interaction of prime type and match was not significant. This pattern provides some indication that the digits in shortcuts, just like the digits in pseudoshortcuts, briefly activated the corresponding number concepts. Apparently, the number concept became suppressed when the meaning of the shortcut was retrieved. This is why there was no match effect for shortcuts at SOA −250.

The inhibition of number concepts in shortcuts appears to be caused by lexical representations, because this is what distinguishes shortcuts from pseudoshortcuts. These lexical representations can be either representations of the words that the shortcuts stand for or specific representations of shortcuts. There is evidence that everyday acronyms, such as STATS, FBI, and WC, have their own phonological codes, consisting of sequences of letter names (Slattery, Pollatsek, & Rayner, 2006), as well as their own lexical representations (Brysbaert et al., 2009). It is therefore likely that shortcuts have their own phonological and lexical representations as well, rather than being merely linked to representations of the words they stand for.

In terms of processing difficulty, our results suggest that embedded digits do not add much to the processing effort of shortcuts. However, the participants in our study were exposed to SMS every day and reported being very familiar with the shortcuts used in the experiment. SMS novices might lack the lexical representations of shortcuts and would therefore have to assemble their phonological forms by retrieving and combining the forms of the individual letters and digits. Under such circumstances, it might be much more difficult to suppress the meanings of embedded digits.

AUTHOR NOTE

REFERENCES

Balota, D. A., Pilotti, M., & Cortese, M. J. (2001). Subjective frequency estimates for 2,938 monosyllabic words. *Memory & Cognition*, **29**, 639-647.

Barber, H. A., & Kutas, M. (2007). Interplay between computational models and cognitive electrophysiology in visual word recognition. *Brain Research Reviews*, **53**, 98-123.

Bowers, J. S., Davis, C. J., & Hanley, D. A. (2005). Automatic semantic activation of embedded words: Is there a "hat" in "that"? *Journal of Memory & Language*, **52**, 131-143.

Brysbaert, M., Speybroeck, S., & Vanderelst, D. (2009). Is there room for the BBC in the mental lexicon? On the recognition of acronyms. *Quarterly Journal of Experimental Psychology*, **62**, 1832-1842.

Carreiras, M., Duñabeitia, J. A., & Perea, M. (2007). READING WORDS, NUMB3R5 and $YMBOL$. *Trends in Cognitive Sciences*, **11**, 454-455.

Coltheart, M., Rastle, K., Perry, C., Langdon, R., & Ziegler, J. (2001). DRC: A dual route cascaded model of visual word recognition and reading aloud. *Psychological Review*, **108**, 204-256.

Dehaene, S., & Cohen, L. (2007). Response to Carreiras et al: The role of visual similarity, feedforward, feedback and lateral pathways in reading. *Trends in Cognitive Sciences*, **11**, 456-457.

Fabre, L., & Lemaire, P. (2005). Age-related differences in automatic stimulus–response associations: Insights from young and older adults' parity judgments. *Psychonomic Bulletin & Review*, **12**, 1100-1105.

Frost, R. (1998). Toward a strong phonological theory of visual word recognition: True issues and false trails. *Psychological Bulletin*, **123**, 71-99.

Ganor-Stern, D., Pinhas, M., & Tzelgov, J. (2009). Comparing two-digit numbers: The importance of being presented together. *Quarterly Journal of Experimental Psychology*, **62**, 444-452.

Libben, G. (1998). Semantic transparency in the processing of compounds: Consequences for representation, processing, and impairment. *Brain & Language*, **61**, 30-44.

Libben, G., Gibson, M., Yoon, Y. B., & Sandra, D. (2003). Compound fracture: The role of semantic transparency and morphological headedness. *Brain & Language*, **84**, 50-64.

Logan, G. D., & Zbrodoff, N. J. (2003). Subitizing and similarity: Toward a pattern-matching theory of enumeration. *Psychonomic Bulletin & Review*, **10**, 676-682.

Nuerk, H.-C., Weger, U., & Willmes, K. (2001). Decade breaks in the mental number line? Putting the tens and units back in different bins. *Cognition*, **82**, B25-B33.

Perea, M., Acha, J., & Carreiras, M. (2009). Eye movements when reading text messaging (txt msgng). *Quarterly Journal of Experimental Psychology*, **62**, 1560-1567.

Perea, M., Duñabeitia, J. A., & Carreiras, M. (2008). R34D1NG W0RD5 W1TH NUMB3R5. *Journal of Experimental Psychology: Human Perception & Performance*, **34**, 237-241.

Pinel, P., Piazza, M., Le Bihan, D., & Dehaene, S. (2004). Distributed and overlapping cerebral representations of number, size, and luminance during comparative judgments. *Neuron*, **41**, 983-993.

Rastle, K., & Brysbaert, M. (2006). Masked phonological priming effects in English: Are they real? Do they matter? *Cognitive Psychology*, **53**, 97-145.

Revkin, S. K., Piazza, M., Izard, V., Cohen, L., & Dehaene, S. (2008). Does subitizing reflect numerical estimation? *Psychological Science*, **19**, 607-614.

Reynvoet, B., Caessens, B., & Brysbaert, M. (2002). Automatic stimulus–response associations may be semantically mediated. *Psychonomic Bulletin & Review*, **9**, 107-112.

Rodd, J. M. (2004). When do leotards get their spots? Semantic activation of lexical neighbors in visual word recognition. *Psychonomic Bulletin & Review*, **11**, 434-439.

Rubinstein, O., Henik, A., Berger, A., & Shahar-Shalev, S. (2002). The development of internal representations of magnitude and their associations with Arabic numerals. *Journal of Experimental Child Psychology*, **81**, 74-92.

Schneider, W., Eschman, A., & Zuccolotto, A. (2002). *E-Prime user's guide*. Pittsburgh: Psychology Software Tools, Inc.

Slattery, T. J., Pollatsek, A., & Rayner, K. (2006). The time course of phonological and orthographic processing of acronyms in reading: Evidence from eye movements. *Psychonomic Bulletin & Review*, **13**, 412-417.

Zhou, X., Chen, C., Chen, L., & Dong, Q. (2008). Holistic or compositional representation of two-digit numbers? Evidence from the distance, magnitude, and SNARC effects in a number-matching task. *Cognition*, **106**, 1525-1536.

Zwitserlood, P. (1994). The role of semantic transparency in the processing and representation of Dutch compounds. *Language & Cognitive Processes*, **9**, 341-368.

APPENDIX
List of Primes Used in the Experiment

Shortcut	Pseudoshortcut	Conventional Spelling
sum1	sud1	someone
t2go	l2go	time to go
g2g	q2g	got to go
wan2	wam2	want to
4u	4o	for you
2nite	2mite	tonight
gr8	qr8	great
in2	iu2	into
1daful	1baful	wonderful
ne1	ni1	anyone
no1	mo1	no one
b4	d4	before
2day	2doy	today
h8	k8	hate
d8	p8	date
w8	v8	wait
ttul8r	ktul8r	talk to you later
l8r	t8r	later
2moro	2noro	tomorrow
l8	t8	late
m8	s8	mate
4ever	4ewer	forever