Orthographic characteristics speed Hindi word naming but slow Urdu naming: evidence from Hindi/Urdu biliterates

Chaitra Rao · Jyotsna Vaid · Narayanan Srinivasan · Hsin-Chin Chen

Abstract Two primed naming experiments tested the orthographic depth hypothesis in skilled biliterate readers of Hindi and Urdu. These languages are very similar on the spoken level but differ greatly in script; Hindi is a highly transparent script, whereas Urdu is more opaque. It was accordingly hypothesized that form-based priming would be greater for Hindi than Urdu, reflecting greater reliance on a phonological assembly route in the more transparent Hindi script. Proficient Hindi/Urdu biliterate readers were presented with primes either in Hindi or Urdu script (Exp. 1), or in Roman transcription (Exp. 2), while targets were always in blocks of Hindi or Urdu. Across both experiments, form-based priming was observed only in Hindi. Additionally, target words were named significantly faster and better in Hindi than in Urdu. The results are taken as support for the hypothesis of differential reliance on phonological assembly as a function of script transparency. Further, the greater graphemic complexity of Urdu script relative to Hindi appears to have contributed to slower and less accurate overall single word reading for Urdu than Hindi, despite the fact that Urdu was the first learned script.

C. Rao (✉)
National Brain Research Centre, Manesar, Haryana 122 050, India
e-mail: chaitlee@gmail.com

J. Vaid
Psychology Department, Texas A&M University, College Station, TX 77843-4235, USA
e-mail: jvaid@tamu.edu

N. Srinivasan
Centre of Behavioural and Cognitive Sciences, Allahabad University, Allahabad, India
e-mail: nsrini@cbcs.ac.in

H.-C. Chen
National Chung Cheng University, Minhsiung, Taiwan
e-mail: hsinchin.chen@gmail.com

Introduction

Many models of visual word recognition assume that skilled readers of any language have two distinct routes by which to access the meaning of a printed word: a direct route which bypasses phonological recoding and an indirect route, which does not (Coltheart, Curtis, Atkins, & Haller 1993; Coltheart, Rastle, Perry, Langdon, & Ziegler 2001; Dijkstra, Grainger, & van Heuven 1999; Perry, Ziegler, & Zorzi 2007). The choice of route may, in turn, be influenced by the relative grapheme-to-phoneme transparency (or shallowness) of a writing system, as proposed in the Orthographic Depth Hypothesis or ODH (Frost, Katz, & Bentin 1987).

Lukatela, Popadić, Ognjenović, and Turvey (1980) showed that readers of Serbo-Croatian (written in Roman and Cyrillic, with both orthographies being highly transparent) were slower at making lexical decisions to words comprised of phonologically ambiguous letters. These words have one rendition if construed as being written in Roman script and another when taken for Cyrillic; for example, POTOP in Roman means *deluge*, whereas in Cyrillic, it is pronounced /rotor/, meaning *rotor*. By contrast, for the more opaque (orthographically deep) Hebrew script, which typically omits vowels, thereby requiring more reliance on prior knowledge or context to infer the correct reading, Bentin, Bargai, and Katz (1984) found that Hebrew readers performed equally well on words with multiple phonological renditions, such as SPR (which may be read /sefer/ —*book*, /sapar/— *barber*, /saper/—*tell*, or /spor/—*count*; phonetic transcriptions taken from Bentin et al., 1984) and those with a single reading, such as KSF (/kesef/, meaning *money*). Further, Frost et al. found that Hebrew readers were significantly faster at making lexical decisions as well as naming aloud words of high than of low frequency, whereas in Serbo-Croatian, both high and low frequency words elicited similar responses across tasks.

Such findings led to the view that readers of shallow orthographies like Serbo-Croatian or Italian rely extensively on a *phonological assembly* route, whereas readers of deep orthographies, such as unvoweled Arabic and Hebrew, predominantly use a *direct access* route in word recognition (Frost et al., 1987; Roman & Pavard, 1987; Tabossi & Laghi, 1992). That is to say, the faithful representation of word phonology in spelling encourages readers of shallow orthographies to convert spelling to sound, which then provides access to meaning. By contrast, deep orthographies have inconsistent or ambiguous spellings, forcing readers to rely on internal, visually-based representations of whole words, which are used to retrieve meaning. These internal representations are thought to be organized with reference to frequency, such that more familiar items are easier to access than less familiar ones.

Although cross-linguistic comparisons led to the formulation of the ODH, factors other than orthographic depth, such as variations in morphology or grammar, might well have contributed to the observed differences in previous cross-linguistic investigations. Consequently, some researchers tried to isolate orthographic depth

effects from effects attributable to other differences across languages. These studies targeted readers of languages such as Japanese and Korean, which make use of different scripts within a single language. Testing phonological priming in Japanese readers, Chen, Yamauchi, Tamaoka, and Vaid (2007) found that recognition of words written using orthographically deep Kanji characters was not facilitated by the previous presentation of homophonic Kanji primes. However, orthographically shallow Hiragana primes effectively facilitated Kanji target recognition. Similar results were also reported by Forster and Yoshimura (cf. Shen & Forster, 1999). Chen et al. interpreted these results to mean that Japanese readers may rely more on a direct access route in processing orthographically deep Kanji, whereas words written in Kana may be accessed more via phonological assembly.

In another study, Korean readers exhibited no effect of word frequency when the stimulus list consisted entirely of words written in the orthographically shallow Hangul script, whereas when stimuli were mostly written in Hanja (which is orthographically much deeper) a significant frequency effect was obtained (Simpson & Kang, 1994). This pattern suggests that in Korean, as in Japanese, readers may preferentially use the direct access route in reading words written in the deep orthography (Hanja or Kanji), whereas words written in the more shallow orthography (Hangul or Kana) may promote the use of phonological assembly.

Researchers have attempted to further isolate the effect of orthographic depth by reducing not only cross-linguistic variation but also intra-linguistic variation, such as differences in word frequency and lexical neighborhood among words of a single language. Thus, Frost (1994) compared lexical decisions and naming using a single set of Hebrew words written in *unpointed* (standard) script as compared to *pointed* script (spelling including vowel diacritics), ensuring that the same words were presented in phonologically opaque (unpointed) versus transparent forms (pointed). When making lexical decisions, the Hebrew readers in Frost's study responded faster to high frequency words, irrespective of whether they were in unpointed or pointed script; however, on the naming task, the effect of word frequency disappeared for words presented in pointed script.

The idea that readers use different strategies to process pointed versus. unpointed script is further supported by the finding that readers were significantly slowed down by the introduction of vowel diacritics in passages of Hebrew and Arabic text (Frost, 1994; Roman & Pavard, 1987). These results suggest that the inclusion of vowels induces readers of these languages to adopt the slower indirect access route. Alternatively, the slower responses noted in the above studies might simply be a byproduct of the relative unfamiliarity of pointed Hebrew and Arabic scripts, rather than reflecting a universal preference for the phonological assembly route when reading a shallow script.

Testing the ODH using Hindi and Urdu

The present study examined the role of orthographic depth in shaping visual word recognition by exploiting the opportunity afforded by the language pair Hindi and Urdu. These two closely related languages belong to the Indo-Aryan language family. They share a common spoken form, often termed Hindustani, which is

believed to have evolved through a commingling of local dialects with Sanskrit, Pali, Prakrit and Persian influences during the Mughal period of Indian history (Kachru, 2008). Notwithstanding their construction as socioculturally distinct languages (Ahmad, 2008) and the slight differences between them in phonemic inventory (Kelkar, 1968), Hindi and Urdu share a common syntactic structure, morphology, and lexicon.

On the other hand, Hindi and Urdu are written in two structurally and visually very different scripts. To illustrate, the Hindustani word for *water*, /pānī/, is पानी in Hindi and پانی in Urdu. Hindi is written and read from left to right whereas Urdu is written and read from right to left. Hindi is written in the highly phonologically transparent, alphasyllabic Devanagari script, whereas Urdu is written in a modified version of the Perso-Arabic script. The basic orthographic unit in Hindi is the *akshara,* which corresponds to a syllable. Consonant aksharas have an implicit schwa vowel associated with them. Except in word-initial positions when they are represented by full aksharas, vowels are represented in the form of diacritics placed nonlinearly above, below, or to the left or right of consonants. Consonant clusters in Hindi may be written using full aksharas, or as ligatures. (See Vaid & Gupta, 2002, for a description of consonant ligatures, and refer to p. 13 of this article for a note on the Hindi schwa deletion rule).

Urdu is an alphabetic script whose letters include thirty-three consonants taken from the Persian alphabet as well as three new letters that represent retroflex sounds and eleven compound letters that represent aspirated consonants (see Mirdehghan, 2010, for a comparative orthographic analysis of Urdu and Persian). Two of the three vowel letters in Urdu also double as consonants—the letter *waau* (و) represents the consonant /ʋ/ in addition to the vowels /o/ and /au/, while the letter *baṛi yey* (ے) stands for the vowel /e/ as well as the consonant /y/. Thus, the Urdu spelling *nuun-waau-peysh* read right to left[1] (پون) may be read /paʋan/ (*wind*) or /paun/ (*three-quarters*). All other vowels in Urdu are represented using vowel diacritics placed above or below the preceding letter. However, standard Urdu script, like Arabic, omits most vowel diacritics, such that many written words have two or more readings. For example, the Urdu spelling *alif-laam-seen* (سلا) may be read /silā/ (*cause to be sewn*) or /sulā/ (*put to sleep*).

In contrast to Urdu, the mapping of spelling to sound is highly consistent in Hindi and vowel marking is obligatory. The omission of vowel diacritics in Urdu, combined with the mapping of multiple phonemes by a single grapheme (and vice versa), means that Urdu is orthographically much deeper compared to Hindi. To illustrate this point, we provide the distinct Hindi forms of the examples above, /paʋan/ (पवन) and /paun/ (पौन), for comparison with the ambiguous Urdu form (پون), and likewise, the Hindi forms of the words /silā/ (सिला) and /sulā/ (सुला) as opposed to the Urdu word form (سلا).

A series of studies by Vaid and colleagues (Vaid, Rao, & Chen 2010) suggests that the difference in orthographic depth between Hindi and Urdu has implications

---

[1] Spellings of Urdu words, where provided, should be read from right to left, in keeping with the directionality of Urdu script. Thus, in the Hindi/Urdu word /paun/, the Urdu letter *peysh* represents /p/, *waau* stands for /au/, and *nuun* is /n/.

for word recognition strategies used by native readers. Evidence from both word naming and lexical decision experiments points to a greater involvement of phonological assembly during skilled reading in Hindi than in Urdu, as word frequency effects were found to be more robust in Urdu, compared to stronger effects of word length in Hindi.

Graphemic complexity as a factor in orthographic depth

Although research on the role of orthography in lexical access has centered almost exclusively on script transparency (orthographic depth), a few studies have examined the influence of *graphemic complexity* in visual word recognition. Aside from being relatively opaque, Arabic orthography is graphemically complex—letters assume different shapes depending on their position within words. For example, the graphemes (ع، ـع and ـعـ) all represent the Arabic letter *ain* (IPA symbol /ʕ/).

In a striking set of studies, Eviatar and colleagues found that native Arabic readers were slower to respond to CVC strings in Arabic as compared to Hebrew and English, their second and third languages respectively (Eviatar & Ibrahim, 2004; Eviatar, Ibrahim, & Ganayim 2004; Ibrahim, Eviatar, & Aharon-Peretz 2002). Indeed, the Arabic readers tested by Eviatar et al. exhibited greater accuracy in identifying Hebrew compared to Arabic letters. Finally, Ibrahim et al.'s study showed bilingual Arabic-Hebrew readers to be faster and more accurate on the visual Trail Making Test (TMT) in Hebrew than in Arabic; the TMT required participants to connect either Arabic or Hebrew letters with Indian numerals in alternating, successive order, from among a random scattering of letters and numbers, similarly to a 'connect the dots' puzzle.

Inasmuch as Urdu is written in a modified Arabic script, it shares the graphemic complexity of Arabic. As in Arabic, Urdu letters have multiple graphemic forms depending on their position within a word. (The Urdu letter *ain* is also represented by ع in isolation, ـع in word-initial and ـعـ in both medial and final positions.) In addition, owing to its cursive property, many graphemes within Urdu words appear highly similar or even identical, and are distinguishable only by the presence and positioning of dots (Mirdehghan, 2010). For example, compare the visual forms of the words /bastī/ (بستی, *village*), /bantī/ (بنتی, *is made*), and /bītī/ (بیتی, *occurred*).

Compared with Urdu, Hindi is much less graphemically complex. Although Hindi has two graphemes each per vowel and consonant, the akshara and ligature graphemes of most Hindi consonants are visually very similar; even diacritics bear some resemblance to their respective vowel aksharas, although less so than consonants. Further, Hindi script is not cursive, making it easier to distinguish individual letters. This contrast in graphemic complexity is evident when comparing the Hindi forms of /bastī/ (बस्ती), /bantī/ (बनती), and /bītī/ (बीती) with the Urdu forms above.

Previous research indicates that graphemic complexity may also differentially influence visual word recognition in Hindi and Urdu. Vaid et al. (2010) noted that skilled readers showed slower lexical decision responses in Urdu as compared to Hindi. The present research sought to test whether a similar pattern would be observed among biliterate readers of Urdu and Hindi, analogous to that observed in Arabic versus Hebrew biliterates.

Form priming and orthographic depth

In the present research, two experiments were conducted to assess primed word naming among skilled, biliterate readers of Hindi and Urdu. The priming technique has been extensively used in word recognition research to study lexical processing in different orthographies. Research shows that a previously exposed word or pseudoword prime that is related either in meaning or in form (phonological and/or orthographic) can facilitate or inhibit the recognition of a subsequently exposed target.

Nevertheless, the time course and nature of priming are influenced by the orthographic depth of a given language. For instance, semantic priming affects processing at the level of whole words but not sub-lexical processing, and therefore, briefly exposed semantic primes are effective in deep orthographies, but have little or no effect on shallow orthographies. Thus, studies on Serbo-Croatian found no effect of semantic primes on word naming (Frost et al., 1987), whereas Hebrew word naming was significantly facilitated by semantic primes (Bentin & Feldman, 1990; Frost et al., 1987).

Conversely, form-matched primes facilitate processing via phonological assembly, and are thus more effective in orthographically shallow languages. For instance, Lukatela, Carello, and Turvey (1990) found that phonological primes facilitated not only word but also pseudoword naming in Serbo-Croatian, and led to faster lexical decisions to real words. Kim and Davis (2002) similarly found robust priming of Korean words written in orthographically shallow Hangul script, both by orthographically dissimilar homophonic primes (analogous to *odd—awed*), and by primes sharing the word onset with the target. Oney, Peter, and Katz (1997) compared word naming in both languages of Turkish-English bilinguals. They found that auditory phonological primes that rhymed with targets exercised a larger facilitative effect on the transparent Turkish orthography compared to the relatively deep English.

Remarkably, Frost and colleagues (Frost, Forster, & Deutsch 1997; Frost, Kugler, Deutsch, & Forster 2005) found that in Hebrew as well as Arabic, form primes (i.e., primes that overlapped with targets in both phonology and orthography) failed to facilitate lexical decision. Thus, word-pairs such as /targil/—/taklit/ (Hebrew: *exercise—record*) and /kamaal/—/Zamaal/ (Arabic: *perfection—beauty*; phonetic transcription follows Frost et al., 2005) did not show priming. Further, Frost et al. (1997) reported an absence of form-based priming even when Hebrew words were named aloud.

Present research

Two experiments with Hindi/Urdu biliterates were conducted. Experiment 1 compared form priming in Hindi versus Urdu. It was expected that word naming in the transparent Hindi orthography would benefit significantly from form-related primes, whereas the opaque Urdu script would show little or no priming. Since form-related primes in Experiment 1 were related in sound and in spelling, Experiment 2 sought to isolate the influence of phonological overlap in form

priming in these languages by presenting primes transcribed in Roman script. As in Experiment 1, it was hypothesized that phonological priming would exert a stronger effect on naming Hindi words than on naming the same words presented in Urdu. Further, in both experiments, the graphemic complexity of Urdu was expected to result in a processing cost, reflected in overall slower responses in Urdu than in Hindi.

By using the language pair Hindi and Urdu, the current study eliminated other potential sources of variability across languages, such as differences due to lexical, morphological, and syntactic variation, as well as eliminating intra-linguistic sources of variation such as differences in lexical frequency, phonological neighborhood, and so on. Additionally, given that our participants were proficient biliterates in Hindi and Urdu, we were able to use a within-subjects design, which allowed us to be more confident that any observed differences in priming could reasonably be attributed to the effect of orthographic depth.

Experiment 1

This experiment assessed the effect of form-related primes on word naming speed for Hindi vs. Urdu target words. Primes, which were presented in the same script as targets, were phonologically monosyllabic CVC words, while targets consisted of bisyllabic words that shared the same initial syllable with their respective primes. (See p. 13 of this article for a note on the Hindi schwa deletion rule.) For example, the prime /hal/ (Hindi: हल; Urdu: حل) meaning *solution*, was paired with the target /halʋā/ (Hindi: हलवा; Urdu: حلوہ) meaning *sweetmeat*. Thus, on a phonological level, there was complete overlap between the prime word and the first syllable of the corresponding target word.

Method

Participants

Eighteen proficient readers of Hindi and Urdu were recruited from a university in north India. They included eight males and ten females in the age range of 18–35 years. All participants claimed Urdu as their native language (in speech and writing) and had been studying Urdu as a major at an advanced level. Participants had all received a minimum of 7 years of formal instruction in Hindi.[2] Their self-rated reading proficiency in both languages averaged 6.5 on a 7-point scale.

Design and materials

A 2 (script: Hindi, Urdu) by 2 (prime type: experimental, control) within-subjects factorial design was used. Stimuli included 256 Hindi/Urdu words, comprising 64

---

[2] Most Hindi/Urdu biliterates tested in both experiments in the present study were multilingual and multiliterate, with knowledge of a diverse array of Indic and non-Indic languages, including Arabic, Bengali, English and Persian.

experimental and 64 control prime-target pairs. Due to the constraints on stimulus selection, it was not possible to control for word class or word frequency.

As outlined above, primes were monosyllabic and targets were bisyllabic. Each experimental prime was identical to the first syllable of the corresponding target, but the prime and target pairs were morphologically and semantically unrelated (e.g., /hal/—/halʋā/). Control prime-target pairs were dissimilar in form as well as being otherwise unrelated; for example, /daur/, meaning *revolution* (Hindi: दौर; Urdu: دور), was paired with /hickī/, *hiccup* (Hindi: हिचकी; Urdu: ہچکی). Additional lists were compiled containing 64 filler pairs and 10 practice pairs, which were similar in length and syllable structure to experimental and control pairs.

In order to maximize the contrast in graphemic complexity between Hindi and Urdu, the CVCCV and CVCCVC targets were carefully selected such that the medial consonants were represented in Hindi by full graphemes rather than by ligatured graphemes. The so-called schwa deletion rule in Hindi requires readers to omit the schwa belonging to word-final consonants, as well as those of non-ligatured medial consonants, leading to less than complete transparency in Hindi spelling. Nevertheless, as illustrated in the introduction section, Hindi spelling achieves considerably greater transparency than Urdu.

Each stimulus was typed in Hindi (Myhindigyan bold font, size 36) and in Urdu (Nastaliq bold, size 68) and presented as individual bitmap images. Whereas similar priming experiments in languages such as English have used upper versus lower case to distinguish primes from targets, we used color in the absence of a case distinction in Hindi and Urdu—primes were presented in yellow and targets in white. Since words written in Urdu script tend to take up more vertical space than their counterparts in Hindi, we attempted to keep individual letter width (as measured in number of pixels) comparable across the scripts.

Procedure

Participants saw 96 prime-target pairs each in separate Hindi and Urdu blocks. Their task was to name the target word aloud as quickly as possible. Within a block, there were 32 items each corresponding to the experimental, control and filler pairs. Assignment of stimuli to blocks was randomized and each prime-target pair was presented in both Hindi and Urdu across different readers. Language order was counterbalanced.

A trial consisted of an initial, central fixation crosshair (800 ms), followed by a forward pattern mask (500 ms). This was replaced by the prime, exposed for either a short (136 ms) or a long (680 ms) duration. A blank screen ISI[3] of 34 ms followed, and then the target appeared and remained on screen until the onset of the participant's response triggered a voice key in the Serial Response Box connected to E-Prime, enabling logging of response latency in milliseconds. Word naming

---

[3] Anticipating that the nature of prime-target overlap might make it difficult for readers to distinguish them as distinct events, we introduced a brief Inter Stimulus Interval (ISI), following the demonstration by Ferrand (1996) that the introduction of an ISI successfully eliminated repetition priming effects.

accuracy was manually coded by the experimenter. ISIs were 1,000 ms long, and a break was allowed after every 48 trials.

Results

Table 1 summarizes participants' mean response latencies and accuracy. Separate 2 (script) × 2 (prime type) repeated measures ANOVA were computed on participants' mean target naming accuracy and reaction time (RT). Outlier trials, including latencies shorter than 250 ms and longer than 2,000 ms were removed preparatory to RT analyses, accounting for 0.2% of Hindi and 1.4% of Urdu trials.

Results of the accuracy analysis showed a main effect of script, reflecting superior accuracy for Hindi compared to Urdu naming, $F(1, 17) = 48.26$, $MSE = 0.003$, $p < .001$. A reliable advantage also emerged for prime type, $F(1, 17) = 10.40$, $MSE = 0.003$, $p = .005$, indicating more accurate naming of words preceded by related than unrelated words. There was no interaction. Analysis of participants' RT also yielded significant main effects of script [$F(1, 17) = 18.62$, $MSE = 970$, $p < .001$] and prime type [$F(1, 17) = 17.41$, $MSE = 977$, $p = .001$], and an interaction of script by prime type, $F(1, 17) = 4.71$, $MSE = 876$, $p < .05$. Simple effects analyses showed a significant difference between form-related and unrelated conditions in Hindi, $F(1, 34) = 20.45$, $MSE = 927$, $p < .01$, but no difference in Urdu ($p > .05$).

Discussion

Consistent with expectation, we found an overall advantage for Hindi over Urdu in naming speed and accuracy, confirming similar findings by Vaid et al. (2010). Moreover, a significant form-based priming effect was observed only in Hindi. The presence of form-based priming in Hindi extends previous findings observed in English and French, among other languages. The lack of form-based priming in Urdu is consistent with a similar finding in Arabic documented by Frost et al. (2005).

One interpretation of the selective form priming in Hindi is that it is an artifact of the greater orthographic overlap between primes and targets in Hindi relative to that in Urdu; the graphemic complexity of Urdu meant that in many prime-target pairs,

Table 1 Mean naming latency and percent accuracy to Hindi/Urdu target words preceded by form-related or unrelated Hindi/Urdu primes, Exp. 1 (n = 18)

	Reaction time (ms)		Accuracy (%)	
	Hindi target words	Urdu target words	Hindi target words	Urdu target words
Same-script form-related primes	706 (35)[a]	821 (45)	98 (1)	90 (1)
Same-script form-unrelated primes	752 (37)	837 (42)	94 (1)	85 (2)

[a] Standard error values in parentheses

the same phoneme was represented by different graphemes, due to their different positions within the word. In order to test this account, data from 16 prime-target pairs in which the Urdu primes were orthographically identical to the first syllable of their respective targets were compared in a separate analysis against the form-unrelated control stimuli. For example, in the pair /šor /—/šorbā/, meaning *hubbub —soup*, the initial syllable of the target retains a visual form identical to the prime: شور—شوربا. Urdu prime-target pairs that did not share an orthographically identical syllable were excluded in this analysis. Nevertheless, no difference was found in Urdu naming speed between targets primed by form-overlapping vs. form-unrelated control words ($MD = 16$ ms, $t < 1$). Thus, the lack of priming observed for Urdu relative to Hindi in our overall analysis is not likely to arise from differences between them in degree of orthographic overlap between primes and targets.

Furthermore, previous research on form priming provides three distinct lines of evidence that also argue against the possibility that our findings for Urdu are attributable to the lower orthographic overlap in the stimuli. Firstly, research on several alphabetic languages, including Dutch, English, French and German, indicates that primes sharing orthographic but not phonological overlap, for instance *axle—able*, tend to inhibit rather than facilitate target recognition, especially when the prime is exposed long enough to be consciously perceived (Davis & Lupker, 2006; de Moor & Brysbaert, 2000; Drews & Zwitserlood, 1995; Ferrand & Grainger, 1993, 1994; Grainger & Ferrand, 1996; Rastle, Davis, Marslen-Wilson, & Tyler, 2000; Seguí & Grainger, 1990). Secondly, studies across languages show that when the degree of orthographic overlap is controlled, primes with greater phonological overlap produce larger facilitation effects; for example, *conal* primes *canal* better than does *cinal* (Drieghe & Brysbaert, 2002; Lukatela & Turvey, 1990; Perfetti & Bell, 1991; Pollatsek, Perea, & Carreiras, 2005).

The third avenue of support comes from research demonstrating that the size of the repetition priming effect is similar across prime-target pairs that change case, irrespective of whether the lower and uppercase letters are visually similar or dissimilar (Bowers, Vigliocco, & Haan, 1998). Bowers et al. assessed identity priming of uppercase targets by lowercase primes among visually similar (*pick— PICK*) as well as dissimilar pairs (*read—READ*), and found no difference in the amount of facilitation in the two conditions, either in word naming or in lexical decision. Thus, the above studies suggest that greater orthographic overlap is not conducive to a larger form priming effect. In fact, primes that resemble targets orthographically are likely to inhibit rather than facilitate recognition.

It would, therefore, appear that the presence of form priming in Hindi but not Urdu in Exp. 1 is attributable to the greater transparency of Hindi script. This conclusion was tested further in Experiment 2, which sought to isolate the phonological component in the priming task. Whereas primes in Experiment 1 were visually as well as phonologically similar to targets, primes in Experiment 2 were only phonologically related to targets. This allowed for a more direct assessment of the contribution of phonological priming during word recognition in Hindi and Urdu.

Experiment 2

This experiment was aimed at comparing phonological priming in Hindi and Urdu. Several steps were taken to ensure that the observed priming would be primarily phonological in nature. First, visual form overlap between primes and targets was eliminated by using Hindi/Urdu words transcribed in Roman script as primes (following the cross-script manipulation used by Chen et al., 2007). For example, the transcription *KHABAR* was used for the prime /xabar/. Additionally, in order to prevent top-down effects resulting from complete lexical activation of the prime, it was decided to exclude one-syllable words as well as words of high frequency. Consequently, primes and targets both were words of two syllables, and fell in the low to medium frequency range (frequencies of 1–250 occurrences per million in the EMILLE/CIIL 2004 Hindi corpus).

The use of Roman-script-transcribed primes that were orthographically less familiar to participants meant that prime processing would be much slower. This consideration, combined with the non-existence in Hindi of homophones analogous to *bare—bear*, which abound in English, necessitated the choice of word-initial rather than word-final (or rhyming) phonological overlap. Primes and targets were therefore phonologically bisyllabic words that had a common first syllable (CV or CVC), but were unrelated otherwise.

It was expected that under these conditions, priming based on visual form similarity would be greatly reduced, allowing for priming based on phonological similarity. It was predicted that phonological priming between initial syllables of primes and targets would be more evident in the relatively transparent orthography of Hindi compared to Urdu.

Method

Participants

Sixteen proficient readers of Hindi/Urdu were recruited on the campus of the same university as in Experiment 1. They included 11 males and 5 females aged 18–30 years, approximately. All except one claimed Urdu as their native language, and all were studying Urdu as a major subject at the University at the time of testing, in addition to having received a minimum of 7 years of formal instruction in Hindi. Participants' self-rated reading proficiency averaged 6.5 on a 7-point scale in both Hindi and Urdu. In addition to Hindi and Urdu, participants had received a minimum of 5 years of formal English (See footnote 1.) instruction, and rated their own English reading proficiency at approximately 5.0 on a 7-point scale.

Design and materials

The experiment used a 2 (script: Hindi, Urdu) × 2 (prime type: phonological, unrelated) within-subjects design. Stimuli included 256 Hindi/Urdu words of two syllables, which made up 128 pairs of words with a common first syllable, for example /xabar/—/xayāl/ (*news—thought*; target in Hindi: ख़याल, target in

Urdu: خیال); primes and targets were morphologically and semantically unrelated. Words within the list were also randomly paired with formally dissimilar (both orthographically and phonologically), semantically unrelated words to create control pairs like /dʒatan/—/xayāl/ (*effort—thought*; Roman prime transcription *JATAN*). Selected pairs included 33 with CVCCVC structure, and 95 with CVCVC structure. As in the first experiment, words with consonant clusters (CC) were chosen only if their Hindi spellings had no ligatures. Due to noun declension and case inflection features of Hindi/Urdu grammar, lemma frequency rather than token frequency was used as the selection criterion, with prime frequency averaging 5.4 occurrences per million or *opm* (SD = 8.7), while mean target frequency was 2.2 opm (SD = 1.6). An additional 32 pairs of words were included as fillers and 10 pairs as practice.

Targets were typed in Hindi (Myhindigyan bold font, size 36) as well as in Urdu (Nastaliq bold, size 68) and presented on a computer screen as individual bitmap images (white text on a black background). Targets across the two scripts were approximately equal in height on the screen. Primes were presented in Courier New lowercase font, size 18 (white text, black background). The spelling used for transcriptions was independently vetted by two judges, one a native Hindi speaker and the other a native speaker of Urdu.

Procedure

A speeded naming task was presented to participants in two blocks of 80 trials each; in one block, targets appeared in Hindi, and in the other in Urdu, with the order counterbalanced across participants. Within a block, participants saw 32 phonological and 32 unrelated prime-target pairs, as well as 16 filler pairs. A trial consisted of an initial, central fixation crosshair (800 ms) followed by the prime word, which was exposed for either a short (136 ms) or a long (680 ms) duration. The target remained on the screen until the participant's voice onset triggered the logging of response latency by E-Prime, as in the first experiment. Naming accuracy was manually coded by the experimenter.

Results

Participants' mean naming latency and accuracy (summarized in Table 2) were compared in separate 2 (script) × 2 (prime type) repeated measures ANOVAs. Reaction time analyses were conducted after eliminating inaccurate responses. Outliers, that is, latencies shorter than 250 ms and longer than 2,000 ms accounted for removal of 0.4% of Hindi trials and 0.8% of Urdu trials.

A significant advantage in accuracy was found for Hindi (98%) versus Urdu (88%) word naming, $F(1, 15) = 54.03$, $p < .001$, but there was no effect of prime type ($F < 1$), nor did script and prime type interact. Analysis of RTs also revealed a main effect of script, with words in Hindi named an average of 134 ms faster than corresponding words in Urdu, $F(1, 15) = 22.84$, $MSE = 12,604$, $p < .001$], but no effect of prime type ($p > .10$). The script by prime type interaction fell short of being significant, with $F(1, 15) = 2.78$, $MSE = 225$, $p = .11$. Planned comparisons

Table 2 Mean naming latency and percent accuracy for Hindi versus Urdu target words preceded by phonologically related or unrelated primes in Roman transcription, Exp. 2 (n = 16)

	Reaction time (ms)		Accuracy (%)	
	Hindi target words	Urdu target words	Hindi target words	Urdu target words
Phonologically related primes	708 (17)[a]	849 (29)	97 (1)	88 (1)
Phonologically unrelated primes	722 (18)	850 (29)	98 (1)	88 (2)

[a] Standard error values in parentheses

using simple effects ANOVA computed separately for Hindi and Urdu RT data revealed that, in Hindi, phonologically primed targets were named marginally faster than words preceded by unrelated primes, $F(1,30) = 3.19$, $MSE = 466$, $p < .05$. In Urdu, the difference between phonologically related and unrelated conditions was not significant ($F < 1$).

Discussion

This experiment confirmed the earlier finding of an overall advantage for Hindi over Urdu in word naming speed. Further, the results revealed different priming patterns in Hindi versus Urdu. Despite the use of visually unfamiliar primes and minimal phonological overlap between primes and targets, Hindi exhibited a tendency for a phonological priming effect, whereas no priming occurred in Urdu.

The facilitative influence of a phonologically identical initial syllable across a Roman-transcribed prime and a Hindi target suggests that readers relied more on the indirect route in identifying Hindi words, wherein the assembly of target phonology was facilitated by prior activation of some of the same phonological units by the prime. Thus, Exp. 2 lent support to the idea that the transparent orthography of Hindi promotes reliance on an indirect, assembly-based route. In contrast, targets presented in the phonologically opaque, graphemically complex Urdu script did not benefit from previous exposure to phonologically related primes.

General discussion

Our aim in this research was to assess the influence of orthographic depth and graphemic complexity of script on word recognition. Two experiments examined primed word naming among highly skilled biliterate readers who were proficient in Hindi and Urdu. In contrast to previous cross-linguistic tests of the ODH, our use of Hindi/Urdu provided a high degree of control over variation in dimensions besides orthography. The common grammar and highly similar morphophonology of Hindi and Urdu enabled us to use stimuli that were identical in sound and meaning across the two languages. Further control was achieved by testing highly proficient, biliterate readers.

Experiment 1 demonstrated that form-related prime-target pairs (comparable to *ham—hamlet*) facilitated word naming in Hindi but not in Urdu. Experiment 2 additionally provided evidence of a small but significant phonological priming effect in Hindi, despite the use of orthographically unfamiliar primes (Roman transcriptions of Hindi/Urdu words) that shared minimal phonological overlap with targets. Neither of the experiments found facilitation of target naming in Urdu. In addition, the results of both experiments showed that participants were faster and more accurate at naming the same words presented in Hindi as compared to Urdu.

When interpreted in the light of previous research, the current findings indicate a dramatic difference in word recognition in Hindi versus Urdu. Earlier studies in English (Feldman & Prostko, 2002; Rastle et al., 2000), French (Seguí & Grainger, 1990), and German (Drews & Zwitserlood, 1995) have consistently shown that form-priming effects dissipate at prime exposure durations exceeding 300 ms. In contrast, Experiment 1 in the present research provides clear evidence of form priming in Hindi, despite the use of relatively long prime exposures (136 and 680 ms). The absence of form priming in Urdu in our data mirrors the lack of form-based priming in Hebrew word naming reported by Frost et al. (1997). However, the Urdu results are singular in light of recent evidence of form priming in both Hebrew (Frost & Yogev, 2001) and Arabic (Boudelaa & Marslen-Wilson, 2005), using very brief prime exposure durations of less than 100 ms.

Equally remarkable is the finding in Experiment 2 of a modest phonological priming effect in Hindi even under conditions that might not be conducive to priming. Unlike a previous study of cross-script priming that employed homophones (Chen et al., 2007), the primes in our Exp. 2 overlapped with targets only in the initial syllable. Therefore, the finding that primes transcribed in Roman facilitated naming of Hindi words further strengthens the argument that processing Hindi script invokes the phonological assembly route to a greater extent (as argued by Vaid et al., 2010).

Taken together, these results strongly suggest that the orthographic differences between Hindi and Urdu play a central role in shaping readers' word recognition strategy. The orthographic shallowness of Hindi appears to encourage reliance on systematic grapheme to phoneme conversion even among highly skilled, adult readers. By contrast, skilled readers of Urdu exhibit little influence of form-based and phonological primes, suggesting predominant reliance on a direct access route in reading Urdu.

Further, both experiments in the current study furnish evidence for a processing cost levied by Urdu orthography—despite being native readers of Urdu, our participants were slower and less accurate at responding to words presented in Urdu than the same words in Hindi script. This pattern is attributable to the greater graphemic complexity of Urdu orthography, and is reminiscent of a similar finding reported for Arabic (Eviatar & Ibrahim, 2004; Ibrahim et al., 2002), whose orthography forms the basis for Urdu script.

Certain limitations qualify the conclusions that may be drawn from the present study. We did not examine the temporal course of word recognition in Hindi and Urdu readers. Previous research with other languages shows that the time course of phonological and orthographic activation follows distinct patterns during visual

word recognition. Lack of form priming in Urdu may have resulted from the coarse-grained manipulation of prime exposure duration[4] in our experiments. However, other research failed to find evidence of form priming in Urdu under highly favorable conditions, that is, using nested targets and a prime-target stimulus onset asynchrony of 50 ms (Rao, 2010).

Studies that manipulate factors such as relative prime and target frequency, and size as well as density of phonological and orthographic neighborhoods of primes and targets are essential for a better understanding of the nature and locus of processing differences between Hindi and Urdu. Further, given the relative scarcity of bilingual word recognition studies on readers of different scripts (e.g., Gollan, Forster, & Frost 1997; Kim & Davis, 2003), it would be interesting in future work to consider cross-script priming effects in Hindi and Urdu readers, that is, where primes are in Hindi and targets in Urdu, or vice versa. Such studies would also help to illuminate the crucial question of whether Hindi and Urdu are indeed cognitively treated as separate languages by readers. Recent evidence indicates that the diglossia prevalent in Arabic, wherein spoken Arabic (SA) differs phonetically, phonologically, morphologically and semantically from literary or modern standard Arabic (MSA) leads readers to treat these as distinct languages (Ibrahim, 2009; Ibrahim & Aharon-Peretz, 2005).

To conclude, although it is still unclear whether Hindi and Urdu are represented and accessed via cognitively distinct lexica, the findings of the current study demonstrate that, at least at the level of script, the contrast in orthographic depth and graphemic complexity exerts clear and dramatic effects on word recognition in Hindi versus Urdu.

References

Ahmad, R. (2008). Scripting a new identity: The battle for Devanagari in nineteenth-century India. *Journal of Pragmatics, 40*, 1163–1183.

Bentin, S., Bargai, N., & Katz, L. (1984). Graphemic and phonemic coding for lexical access: Evidence from Hebrew. *Journal of Experimental Psychology. Learning, Memory, and Cognition, 10*, 353–368.

Bentin, S., & Feldman, L. B. (1990). The contribution of morphological and semantic relatedness to repetition priming at short and long lags: Evidence from Hebrew. *Quarterly Journal of Experimental Psychology, 42A*, 693–711.

Boudelaa, S., & Marslen-Wilson, W. D. (2005). Discontinuous morphology in time: Incremental masked priming in Arabic. *Language and Cognitive Processes, 20*, 207–260.

---

[4] Prime exposure was varied at short (136 ms) and long (680 ms) SOA in both experiments, but preliminary analyses revealed only a main effect of SOA on participants' naming latencies, with faster responses at the long SOA, and no interactions with other factors. Therefore, SOA was dropped from subsequent analyses.

Bowers, J. S., Vigliocco, G., & Haan, R. (1998). Orthographic, phonological, and articulatory contributions to masked letter and word priming. *Journal of Experimental Psychology: Human Perception and Performance, 24*, 1705–1719.

Chen, H. C., Yamauchi, T., Tamaoka, K., & Vaid, J. (2007). Homophonic and semantic priming of Japanese Kanji words: A time course study. *Psychonomic Bulletin & Review, 14*, 64–69.

Coltheart, M., Curtis, B., Atkins, P., & Haller, M. (1993). Models of reading aloud: Dual-route and parallel-distributed-processing approaches. *Psychological Review, 100*, 589–608.

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

Davis, C. J., & Lupker, S. J. (2006). Masked inhibitory priming in English: Evidence for lexical inhibition. *Journal of Experimental Psychology: Human Perception and Performance, 32*, 668–687.

De Moor, W., & Brysbaert, M. (2000). Neighborhood-frequency effects when primes and targets are of different lengths. *Psychological Research, 63*, 159–162.

Dijkstra, A. F. J., Grainger, J., & van Heuven, W. J. B. (1999). Recognition of cognates and interlingual homographs: The neglected role of phonology. *Journal of Memory and Language, 41*, 496–518.

Drews, E., & Zwitserlood, P. (1995). Orthographic and morphological similarity in visual word recognition. *Journal of Experimental Psychology: Human Perception and Performance, 21*, 1098–1116.

Drieghe, D., & Brysbaert, M. (2002). Strategic effects in associative priming with words, homophones, and pseudohomophones. *Journal of Experimental Psychology. Learning, Memory, and Cognition, 28*, 951–961.

EMILLE/CIIL. (2004). *Corpus of South Asian Languages*. European Language Resources Association. Retrieved from http://catalog.elra.info/product_info.php?products_id=696.

Eviatar, Z., & Ibrahim, R. (2004). Morphological and orthographic effects on hemispheric processing of nonwords: A cross-linguistic comparison. *Reading and Writing: An Interdisciplinary Journal, 17*, 691–705.

Eviatar, Z., Ibrahim, R., & Ganayim, D. (2004). Orthography and the hemispheres: Visual and linguistic aspects of letter processing. *Neuropsychology, 18*, 174–184.

Feldman, L. B., & Prostko, B. (2002). Graded aspects of morphological processing: Task and processing time. *Brain and Language, 81*, 1–16.

Ferrand, L., & Grainger, J. (1993). The time course of orthographic and phonological activation in the early phases of visual word recognition. *Bulletin of the Psychonomic Society, 31*, 119–122.

Ferrand, L., & Grainger, J. (1994). Effects of orthography are independent of phonology in masked form priming. *Quarterly Journal of Experimental Psychology, 47A*, 365–382.

Frost, R. (1994). Prelexical and postlexical strategies in reading: Evidence from a deep and a shallow orthography. *Journal of Experimental Psychology. Learning, Memory, and Cognition, 20*, 1–16.

Frost, R., Forster, K. I., & Deutsch, A. (1997). What can we learn from the morphology of Hebrew? A masked priming investigation of morphological representation. *Journal of Experimental Psychology. Learning, Memory, and Cognition, 23*, 829–856.

Frost, R., Katz, L., & Bentin, S. (1987). Strategies for visual word recognition and orthographical depth: A multilingual comparison. *Journal of Experimental Psychology: Human Perception and Performance, 13*, 104–115.

Frost, R., Kugler, T., Deutsch, A., & Forster, K. I. (2005). Orthographic structure versus morphological structure: Principles of lexical organization in a given language. *Journal of Experimental Psychology. Learning, Memory, and Cognition, 31*, 1293–1326.

Frost, R., & Yogev, O. (2001). Orthographic and phonological computation in visual word recognition: Evidence from backward masking in Hebrew. *Psychonomic Bulletin & Review, 8*, 524–530.

Gollan, T., Forster, K. I., & Frost, R. (1997). Translation priming with different scripts: Masked priming with cognates and noncognates in Hebrew-English bilinguals. *Journal of Experimental Psychology. Learning, Memory, and Cognition, 23*, 1122–1139.

Grainger, J., & Ferrand, L. (1996). Masked orthographic and phonological priming in visual word recognition and naming: Cross-task comparisons. *Journal of Memory and Language, 35*, 623–647.

Ibrahim, R. (2009). The cognitive basis of diglossia in Arabic: Evidence from a repetition priming study. *Psychology Research and Behavior Management, 2*, 93–105.

Ibrahim, R., & Aharon-Peretz, J. (2005). Is literary Arabic a second language for native Arab speakers?: Evidence from semantic priming study. *Journal of Psycholinguistic Research, 34*, 51–70.

Ibrahim, R., Eviatar, Z., & Aharon-Peretz, J. (2002). The characteristics of Arabic orthography slow its processing. *Neuropsychology, 16*, 322–326.

Kachru, Y. (2008). Hindi-Urdu-Hindustani. In B. B. Kachru, Y. Kachru, & S. N. Sridhar (Eds.), *Language in South Asia* (pp. 81–112). Cambridge, England: Cambridge University Press.

Kelkar, A. (1968). *Studies in Hindi–Urdu I: Introduction and word phonology*. Poona: Deccan College.

Kim, J., & Davis, C. (2002). Using Korean to investigate phonological priming effects without the influence of orthography. *Language and Cognitive Processes, 17*, 569–591.

Kim, J., & Davis, C. (2003). Task effects in masked cross-script translation and phonological priming. *Journal of Memory and Language, 49*, 484–499.

Lukatela, G., Carello, C., & Turvey, M. T. (1990). Phonemic priming with words and pseudowords. *European Journal of Cognitive Psychology, 2*, 375–394.

Lukatela, G., Popadić, D., Ognjenović, P., & Turvey, M. T. (1980). Lexical decision in a phonologically shallow orthography. *Memory & Cognition, 8*, 124–132.

Lukatela, G., & Turvey, M. T. (1990). Automatic and pre-lexical computation of phonology in visual word identification. *European Journal of Cognitive Psychology, 2*, 325–343.

Mirdehghan, M. (2010). Persian, Urdu, and Pashto: A comparative orthographic analysis. *Writing Systems Research, 2*(1) (in press).

Oney, B., Peter, M., & Katz, L. (1997). Phonological processing in printed word recognition: Effects of age and writing system. *Scientific Studies of Reading, 1*, 65–83.

Perfetti, C. A., & Bell, L. (1991). Phonemic activation during the first 40 ms of word identification: Evidence from backward masking and priming. *Journal of Memory and Language, 30*, 473–485.

Perry, C., Ziegler, J. C., & Zorzi, M. (2007). Nested incremental modeling in the development of computational theories: The CDP+ model of reading aloud. *Psychological Review, 114*, 273–315.

Pollatsek, A., Perea, M., & Carreiras, M. (2005). Does conal prime CANAL more than cinal? Masked phonological priming effects in Spanish with the lexical decision task. *Memory & Cognition, 33*, 557–565.

Rao, C. (2010). *Morphology in word recognition: Hindi and Urdu*. Unpublished doctoral dissertation, Department of Psychology, Texas A&M University.

Rastle, K., Davis, M. H., Marslen-Wilson, W. D., & Tyler, L. K. (2000). Morphological and semantic effects in visual word recognition: A time course study. *Language and Cognitive Processes, 15*, 507–537.

Roman, G., & Pavard, B. (1987). A comparative study: How we read Arabic and French. In J. K. O'Regan & A. Levy-Schoen (Eds.), *Eye movements: From physiology to cognition* (pp. 431–440). Amsterdam: North Holland Elsevier.

Seguí, J., & Grainger, J. (1990). Priming word recognition with orthographic neighbors: Effects of relative prime-target frequency. *Journal of Experimental Psychology: Human Perception and Performance, 16*, 65–76.

Shen, D., & Forster, K. I. (1999). Masked phonological priming in reading Chinese words depends on the task. *Language and Cognitive Processes, 14*, 429–459.

Simpson, G. B., & Kang, H. W. (1994). The flexible use of phonological information in word recognition in Korean. *Journal of Memory and Language, 33*, 319–331.

Tabossi, P., & Laghi, L. (1992). Semantic priming in the pronunciation of words in two writing systems: Italian and English. *Memory & Cognition, 20*, 315–328.

Vaid, J., & Gupta, A. (2002). Exploring word recognition in a semi-alphabetic script: The case of Devanagari. *Brain and Language, 81*, 679–690.

Vaid, J., Rao, C., & Chen, H.-C. (2010). The role of script: Word recognition in Hindi vs. Urdu. Unpublished manuscript, Texas A&M University.