ELUSIVE ASPECTS OF VISUAL WORD FORM PROCESSING: ROTATIONS AND DIACRITICS

András Benyhe MD

PhD Thesis

Szeged

2023

University of Szeged

Albert Szent-Györgyi Medical School

Doctoral School of Theoretical Medicine

Elusive Aspects of Visual Word Form Processing: *Rotations* and Diäcritics

PhD Thesis

András Benyhe MD

Department of Physiology, Albert Szent-Györgyi Medical School

University of Szeged, Szeged

Supervisors:

Péter Csibri PhD

Institute of Psychology, Faculty of Arts, University of Szeged, Szeged

Gyula Sáry MD PhD DSc

Department of Physiology, Albert Szent-Györgyi Medical School

University of Szeged, Szeged

Szeged

2023

Table of contents

List of publications

I. **Benyhe, András**; Csibri, Péter

"Can rotated words be processed automatically? Evidence from rotated repetition priming"

MEMORY & COGNITION, 49: 1163-1171 (2021)

IF: 2.481; SJR indicator (Experimental and Cognitive Psychology): Q1

https://doi.org/10.3758/s13421-021-01147-4

II. **Benyhe, András**; Labusch, Melanie; Perea, Manuel

"Just a mark: Diacritic function does not play a role in the early stages of visual word recognition"

PSYCHONOMIC BULLETIN & REVIEW (2023)

IF_{2021}: 4.412; SJR indicator (Experimental and Cognitive Psychology): D1

https://doi.org/10.3758/s13423-022-02244-4

Cumulative IF: 6.893

List of abbreviations

BOLD, blood-oxygen-level-dependent

CI, confidence interval

CrI, credible interval

EMM, estimated marginal mean

ERP, event-related potential

HNC, Hungarian National Corpus

IT, inferior temporal

LCD, local combination detector

GLMM, generalized linear mixed-effects model

SERIOL, Serial Encoding Regulated by Inputs to Oscillations within Letter units

RT, response time

VWFA, visual word form area

1 Introduction

In recent centuries, reading and writing has become a critically important mode of communication among humans; it allows us to code and conserve linguistic information, and transport it through space and time while maintaining its original form. Literacy is undoubtedly an important driver of recent cultural evolution, capable of changing the way we think (Mullins et al., 2013; Pinheiro et al., 2020). Yet, the time window since the development of writing (and even more so since the widespread extent of literacy), is considered to be too short for substantial biological evolution, meaning that we are bound to read and write with an ancient, pre-literate brain (Dehaene & Cohen, 2007). Apparently, “one can teach an old brain new tricks”. And crucially, reading-specific functions seem to inhabit the same cortical area in all humans, the visual word form area (VWFA) in the left fusiform gyrus (McCandliss et al., 2003). Figuring out how this system works could allow us to understand how it emerges, and how we may approach its malfunctions. In my thesis, I summarize my behavioral research on visual word form processing and reflect on its implications in understanding the underlying neural system, focusing on two aspects that commonly eluded major theories of word identification: the effect of word orientation and the nature of diacritical letters.

1.1 The visual system

We understand the world through our perceptions, the richest of which is arguably the visual modality. Light rays of various energies are hitting the environment, and they are absorbed, reflected or scattered differentially, depending on the properties of the object they hit. Being able to capture this light has proven to be advantageous through the eons, guiding the evolution of a most peculiar organ, the eye. From the simple act of tracking daylight cycles, through the ability to sense directional light and regulate behavioral states, the modern eye has taken a shape that is capable of a high-resolution reconstruction of the environment and guides complex interactions and even cognitive functions (Nilsson, 2021). The precise optics of the human eye allow for the preservation of visual information on the retina, where various aspects of the stimuli are detected and coded. The optic nerve then transmits this information to the brain at an estimated rate of 8 Mbits/s (Koch et al., 2006), a number that roughly equals an HD (1080p) video with a 24 Hz frame rate. This data stream is relayed to the visual cortex, where the information is organized, combined, and decoded. Through such grouping of visual information, the mind can build a rich model of the visual environment, the benefit of which obviously includes foraging, hunting or evasion behaviors, all clearly linked to survival and fitness. In our modern human lives, the detection and decoding of squiggly lines on paper can also prove to be useful but in other ways. By reading and writing, we uphold our complex society (Mullins et al., 2013): we can learn in unified ways, we can keep records of the things we own, we can remember events and people from the past, or testify love to our significant others (Figure 1).

Figure 1 The message on this wall is banal, and yet deeply human. An attempt at immortality akin to those at Lascaux cave but with a completely transparent meaning ("My Baby!! I Will Always LOVE You!!"); an absurd tribute to human ingenuity and the complexities of our perception.

Importantly, the line of processing the visual system takes is not a serial 'pixel by pixel' analysis, and thus the bandwidth comparison from the last paragraph is somewhat misleading. Neural circuitry already in the retina allows for compression of the raw input of the photoreceptors towards the output of the ganglion cells. Light adaptation at multiple levels allows for a wider dynamic range than what the firing rate range of ganglion cell axons would allow for, and sideways connections allow for comparison of illumination between the center and the surround of receptive fields, thus filtering the raw photoreceptor activations. This means that the lateral geniculate nucleus (the thalamic nucleus devoted to relaying the retinal output to the visual cortex), already represents the visual information as local contrasts, i.e., brightness and color values are relative and not absolute.

The cortical network dedicated to analyzing the shape and identity of objects lies along the ventral visual stream, a series of interconnected areas spanning from the occipital pole through occipitotemporal areas to the inferior temporal (IT) cortex. The structure of this network is based on a serial bottom-up hierarchy, where subsequent levels are occupied with larger and larger portions of the visual field and are responsive to increasingly complex aspects of the stimuli. At each level, the information is coded by the firing rate in the population of active neurons, and is passed on in a feedforward manner (DiCarlo et al., 2012). The neural units of the same level work relatively independently from each other, in a parallel fashion, so many possible combinations of their activity are possible. The summed neural activity across a subset of neurons forms the input of the neurons at the next level. At the end of the line, the response pattern of the neuronal ensemble at the IT level is sufficient to account for object recognition (Hung et al., 2005).

Although the evoked potential latencies clearly point to the validity of the feedforward hierarchy in the ventral stream (Schmolesky et al., 1998), the existence of feedback connections is also necessary to account for phenomena observed in low signal-to-noise ratio scenarios (e.g., captchas, where human visual perception outperforms feedforward convolutional networks) and perceptual learning (Ahissar & Hochstein, 2004; Kravitz et al., 2013). Such top-down modulation would support targeted search, or even be the basis of back-propagation during training.

Much of the fine details that we know about visual processing comes from non-human primate experiments, including specific tasks such as body and face recognition (Vogels, 2022). For the purposes of the topic of this thesis, this poses a serious problem: only humans can read (although, other primates were shown to possess the visual capacities required for word recognition: Grainger et al., 2012; Rajalingham et al., 2020). In the following, I would like to review the current understanding on how the visual system serves orthographical analysis, the act of recovering the meaning of a word from a cluster of written symbols.

1.2 Neural background of reading

Experienced readers (likely exemplified by the reader of this text) are able to decipher written words at a glance and with apparent ease. The eyes seem to glide smoothly along the lines, but in reality, the eyes move in quick saccades and stop for short fixations on words. It is only during these fixations that visual processing can take place, while the letters stay still on the fovea. In the 250 ms that fixations take on average, we can recognize letters and their relative order to identify a word, and even get some initial information about the following words in the line (Rayner, 1998).

Importantly, the decoding of all letters in a word must happen simultaneously, in a parallel fashion, and then pieced back together quickly to explain how word identification is possible in a single glance. As we have seen, the capability of such parallel processing is present in the ventral visual stream, and the earliest of evidence has already pointed to the involvement of these areas in reading (Dejerine, 1892). The affected area that caused ‘word blindness’ in the case reports of Dejerine (the occipitotemporal sulcus and fusiform gyrus) was later identified as the VWFA (Cohen et al., 2002; McCandliss et al., 2003), and its integration into the ventral visual stream is of key importance in the most influential neural model of reading, the local combination detector (LCD) model (Dehaene et al., 2005).

The LCD model offers a robust, hierarchical representation of visual word form processing. It builds on the convergence pattern seen along the ventral visual pathway: co-occurrence of local lower-level features can be combined into higher level features (Figure 2). Starting from the well-known phenomenon of orientation-sensitivity of V1 arising from multiple point like inputs arranged in a row (Hubel & Wiesel, 1962), we can follow the same idea to the emergence of invariant letter representations (Dehaene et al., 2005). The invariance is a crucial step, as letters come in many shapes and sizes, and, due to the limitations of neural resources, need to be coded abstractly. Thus, 'THIS' is understood as the same word as 'this' or even '*this*', despite the obvious differences in visual appearance. The emergence of invariance is not too unexpected, since letters are built from the combination of a finite number of simple features. So, it is not the exact shape of the letter that needs to be decoded, just the set of its defining features. Contours meeting end-to-end as in 'L' or end-to-side as in 'T', or simply crossing, as in 'X' are very common features in natural environments too, so it is very likely, that writing systems have converged towards simplicity to decrease computational load and to facilitate acquisition (Changizi et al., 2006). Once the letters of a word are encoded, the word form processing system must also decode their order, a crucial step to differentiate anagrams (e.g., 'night' vs. 'thing'). A serial code that is based on absolute letter position in a word would be very fragile and error prone, as the code would easily shift. To account for this, Dehaene et al. (2005) apply the theory behind the SERIOL model (Serial Encoding Regulated by Inputs to Oscillations within Letter units; Whitney, 2001): a word representation based on open bigrams, letter pairs, whose relative positions are known. The nodes of this level would only be activated when both letter units are active and are positioned in the right order, and they can even have a few characters in between, hence the name 'open'.

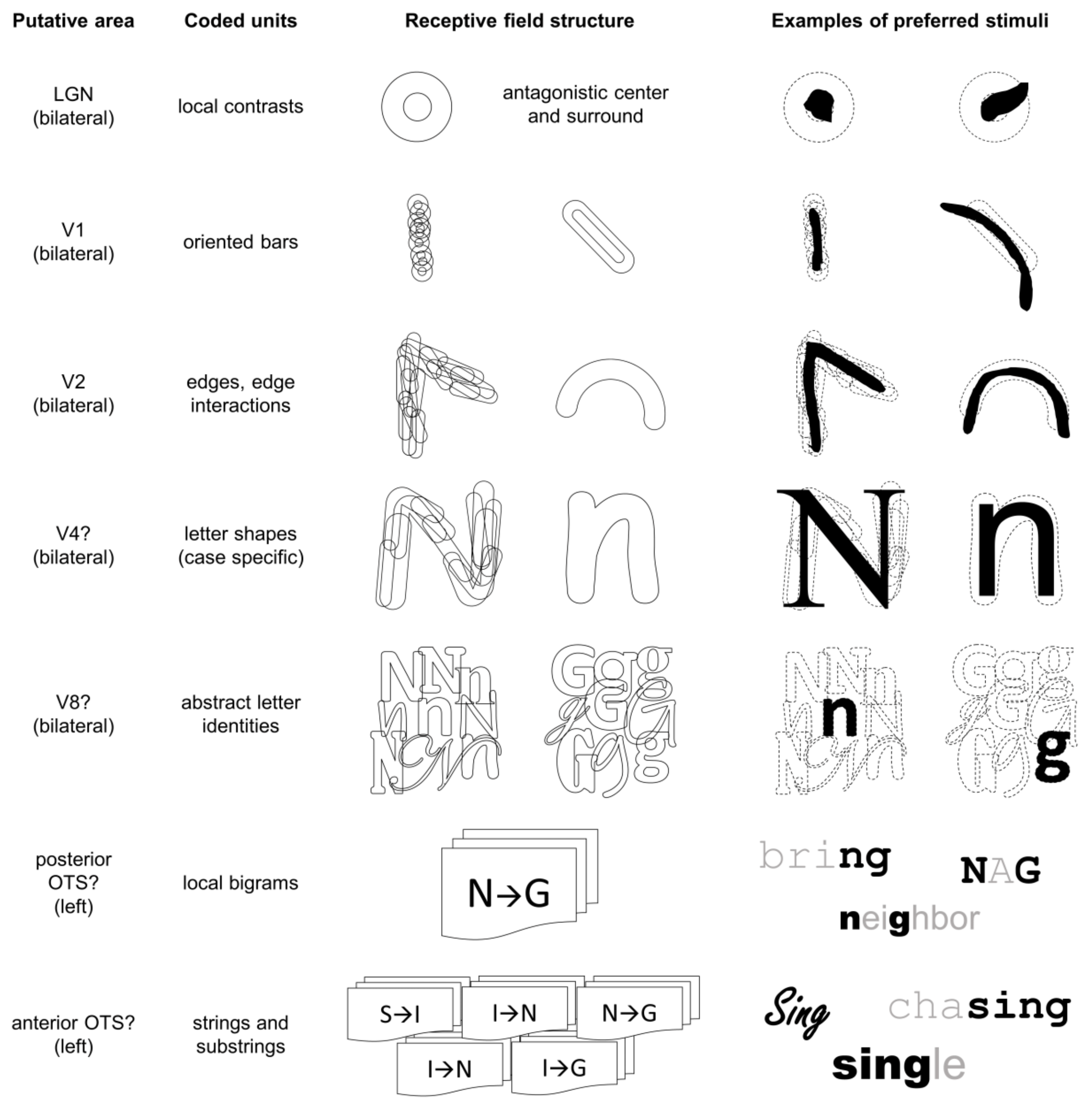


Figure 2 The local combination detector (LCD) model of visual word form processing and word recognition. Subsequent processing levels receive convergent information from previous ones, thereby increasing receptive field size and complexity while also gaining invariance. The featured areas are the lateral geniculate nucleus (LGN) and the primary and secondary visual cortices (V1 and V2), whose properties are well documented. Further extrastriate visual areas (V4 and V8) and inferotemporal areas such as the occipitotemporal sulcus (OTS) are hypothesized to accommodate the higher processing functions of the model. Based on Dehaene et al. (2005).

Although it is understood that the VWFA has broader functionality than just visual word form processing (Price & Devlin, 2003; Vogel et al., 2014), it is still a crucial point in the line of processing during word recognition. Clearly, the VWFA could not have evolved for reading, due to its recency on an evolutional timescale; reading first appeared around 5000 years ago but became widespread only after the invention of the printing press in the 15th century, and the education of the masses after the industrial revolution. The controversy between our ancient brain and its recent function is addressed by the neural recycling hypothesis (Dehaene & Cohen, 2007): novel functions must find their neuronal niche, a region whose organization and connections make it suitable to host the specific function. Since the original functional organization is given, it puts strong constraints on the acquisition and capabilities of the novel function, however some rewiring is also possible through neuronal plasticity. Connectivity studies have found, that the VWFA has extensive connections not only with nearby visual areas but also with language networks, making it ideal for the visual to linguistic mapping (Bouhali et al., 2014; Stevens et al., 2017). The curious finding of VWFA activation in congenitally blind Braille readers, however, ask for a more nuanced theory (Büchel et al., 1998; Reich et al., 2011). Indeed, it is possible, that the VWFA is not really visual, but rather a multimodal area, linking spatial attention and language functions (Chen et al., 2019). Therefore, it is likely, that the VWFA aids in pointing to specific sensory inputs to be decoded as linguistic information, as a way of attention modulation. Another interesting finding in support of the idea that the VWFA acts merely as an information server for linguistic networks is that the rate of saccades during natural reading matches the rate of speech production (Gagl et al., 2022). This, however, does not change the interpretation that during reading, the VWFA is the interface between highly structured visual information and its meaning.

1.3 Computational models of visual word form processing

The localization of orthographic processing in the brain is apparently solved, yet the computations that take place within are still of debate. Although the neural substrates of reading can be reached by non-invasive manipulations (Arrington et al., 2023), these cannot achieve high enough precision to target specific steps of orthographic processing. We are therefore bound to develop theoretical models that can explain the phenomena observed in humans (e.g., word-frequency effect, word-superiority effect, letter transposition effects, etc.; Reichle, 2021). Most popular computational models are of the connectionist type. Essentially, they are mathematical functions – or rather, systems of functions – often in the form of artificial neural networks mimicking the properties of real neuron populations. The units of the network are organized in layers, usually an input, an output and one or more hidden (deep) layers are needed. The activation (roughly representing firing rate) of the input layer propagates through connections to the subsequent layers, where the resulting activity depends on the weights of its connections. The models can usually learn in a controlled way, by tuning the unknown variables (e.g., the weights of the connections) so that the outputs have the least errors. After the model is trained, it can predict the outputs of various tasks, similar to humans in experimental conditions. Computational models are of exceptional use, as their predictions are comparable to experimental data, and they are ideally falsifiable through novel experimental findings.

Regardless of writing system, the universal task in written word recognition (and also in word recognition modeling) is the mapping of visual objects to linguistic units (Li et al., 2022). In alphabetic systems (e.g., current European written languages) each writing symbol or group of symbols represents a phoneme or a group of phonemes (i.e., the written word directly represents the sounds of the spoken word), thus phonology might also play a part in the decoding process. Many influential models, such as the triangle model (Seidenberg & McClelland, 1989) or the dual-route cascaded model (Coltheart et al., 2001) include two routes from script to meaning: a direct route from orthography to semantics, and an indirect route with phonology in between, connected at multiple levels of processing. Importantly, the indirect route has to account for inconsistencies in grapheme-to-phoneme associations, the relevance of which depends on orthographic depth. Deep orthographies, such as English, have many rules for pronunciation and spelling, whereas shallow orthographies, like Hungarian, have a more direct, almost one-to-one grapheme-phoneme correspondence (Borgwaldt et al., 2005). Readers of deep orthographies necessarily rely more on the direct route, and read words as images, whereas the indirect route is more readily available in the case of shallow orthographies.

Although the above-mentioned models explain some important phenomena of reading (naming, lexical decision, word-frequency effect, word-superiority effect), they come with the assumption that letters of a word are already encoded in fixed slots at the input stage. In practice, this means, that their input translates into something like “letter *r* in position 1, letter *e* in position 2, letter *a* in position 3, and letter *d* in position 4”, making up the word *read*. The problem here is that the pseudowords *raed* and *riod* would perform equally under this assumption, since they both have inaccurate letters in positions 2 and 3. In human reading however, we find that letter-transpositions (*raed*) are easily overlooked, whereas letter substitutions (*riod*) are more easily rejected (Grainger & Whitney, 2004). Aiming at this deficiency, multiple models were developed to resolve the problem of letter positions. The SERIOL model (Whitney, 2001) shows that the side of the visual field and the relative visual acuity at which a letter is perceived, can be used to localize each letter of a string along a positional gradient ordered from left to right. According to the model, the value along this gradient defines the response characteristics of the letter detectors (firing delay within a latent oscillatory cycle), which in turn can be compared to each other. This gives rise to the open bigram detectors (as applied in the LCD model, Figure 2; Dehaene et al., 2005), and creates a distributed code containing relative positions of letter pairs (e.g., the word *read* has the following bigrams: *re, ra, rd, ea, ed, ad*). A letter transposition in this case would affect only a subset of the bigrams, damaging the code much less than in a slot-based model. Another model focusing on letter positions is the overlap model (Gomez et al., 2008), which explains letter transposition effects particularly well. This model assumes that there is uncertainty about letter positions – especially when a word is only briefly presented – and the letter position probabilities overlap. The positions within the code are estimated from the perceptual normal distributions.

Another model worth mentioning here would be the Bayesian Reader model (Norris, 2006), which formulates reading in a Bayesian framework. Its assumption about the orthographic code is once again slot-based, but at each position, every letter of the alphabet can have some probability. The letters with the greatest probability can be read out into a decoding framework. During observation, information is initially scarce, and the output is based mostly on prior experience, but with time information builds and the output is increasingly certain (converges onto a code containing letters with probabilities close to one). This can explain uncertainties in letter identity (e.g., O is similar to Q) and in letter position as well (a perceived letter increases the likelihood not only in its own position but in neighboring positions as well).

1.4 Missing components of modeling

1.4.1 Rotations

In our modern world full of writing, it is very common to come across words that are rotated in some way: titles on the spines of books are usually rotated in 90°; when multiple guests put their heads together to read a single restaurant menu, someone will always be bound to read it in a suboptimal angle; the scrabble board is upside down from the other side of the table; the television does not rotate with us when we watch it lying on our side, yet we go on reading the subtitles. Likewise, during handwriting, many individuals rotate the paper in some degree (especially left-handers to neutralize wrist position), thus forming a strong personal experience with rotated texts. Although there is evidence that in-plane rotation has a serious cost in the reading process (Koriat & Norman, 1985), computational models rarely account for this effect (but see, Whitney, 2002). Particularly, the question is, how the brain solves the position problem when the letters are not in the usual horizontal, left-to-right alignment.

Cohen et al. (2008) have found that whenever words were observed in a degraded manner (e.g., rotated or with increased letter spacing), response time (RT) increased steeply in a nonlinear fashion, accompanied by an increased blood-oxygen-level-dependent (BOLD) signal both in the ventral and dorsal visual pathways. They interpreted these results in terms of the LCD model (Dehaene et al., 2005), stating, that degradation above a certain threshold makes the VWFA no longer capable of decoding words, and it requires auxiliary mechanisms from the dorsal stream. Like Whitney (2002), Dehaene et al. also suggest that this is done by mental rotation, the act of imagining how the word (or its constituent letters) would look like counter-rotated and doing the orthographical analysis afterwards. They identified this blockade at the letter level, as BOLD signals correlated with response latencies slightly posterior to the peak WVFA activation to words. They conclude that rotation above 45° interrupts letter processing, which in turn cannot feed into the quick, parallel decoding of the word, and thus the brain must rely on slower, serial mechanisms aided by the dorsal stream.

Despite some earlier results suggesting that letter rotation disrupts letter recognition (Risko et al., 2014), Perea et al. (2018) found strong evidence for the opposite in a masked identity priming letter matching experiment. In their study, the size of the unconscious priming effect did not depend on the rotation of the briefly presented prime, showing more rotation resistant processing than expected. In another experiment, Perea et al. (2020) have also shown that the orientation of the letters did not matter in the context of words either, when they compared marquee (words formed by normally oriented letters stacked vertically and read from the top) and 90° rotated words in a masked priming lexical decision experiment. This also points to the remarkable resilience of letter detectors in the reading network and questions the validity of the 45° barrier stated by Cohen et al. (2008).

To resolve this contradiction, we might look to other fields of visual perception. For example, in the case of object recognition, evidence was found for isolated object orientation agnosia: a spared ability to recognize rotated objects with the inability to tell their orientation (Harris et al., 2001). Indeed, it is logical to argue that an object needs to be identified, before we can tell its orientation, so identity comes first. The orientation invariance of object recognition was proven in repetition blindness (Harris & Dux, 2005) and repetition priming paradigms as well (Harris et al., 2008). Interestingly, when tested with letter shapes, the repetition blindness effect was even greater when the repeated letter appeared in different orientations (Corballis & Armstrong, 2007), again highlighting that letter detectors have orientation invariance, and that the cost of rotation in reading must originate from later processing steps.

In a simulation, Hannagan et al. (2021) have found that a convolutional artificial neural network, mimicking the ventral visual stream, could easily be trained for word recognition. This study serves interesting evidence for the neural recycling hypothesis, as the network was originally trained on objects, and words were only introduced later. Crucially, it reproduced many of the characteristics of human word recognition, like letter transposition effects, or alexia caused by acquired lesions. The rotation tolerance of the model was however only 10°, which was explained by the fact that the training dataset contained only normally oriented words. In this case, it might just be true that the letter detectors failed when presented with rotated inputs, but for humans exposed to rotated letters all the time (thus aiding perceptual learning), processing would be expected to go beyond this step.

In summary, the current notion on what happens in the brain when observing rotated texts is controversial. In order to build models capable of solving the problem of rotations and serve powerful predictions in a host of reading tasks, first we need to better explore the phenomena involving rotated words.

1.4.2 Diacritics

A large number of written languages are based on the Latin alphabet, even though it is very unsuitable to meet the requirements of each language in capturing the relevant nuance in pronunciation. Many writing systems coped with this shortcoming by making amends to the alphabet to better meet their needs. The most common modification is the addition of diacritics, small markings usually above a letter, which change the linguistic information of the given character. Importantly, the linguistic function of diacritics can vary among languages. English readers might only come across the occasional diaeresis, two dots above a letter, indicating that it has to be pronounced separately and not as part of a diphthong (e.g., in coöperate, although it is considered archaic). In Spanish, the usage of diacritics is more frequent, and the acute accent mark indicates the location of lexical stress within a word but does not change vowel sound quality. In the case of Hungarian, diacritics are used to distinguish the fourteen unique vowels of the spoken language by modifying the five Latin base vowels. Then again, in French there is a wide variety of diacritics for a host of reasons, some etymological (e.g., the circumflex in ‘forêt’ indicates the historical deletion of ‘s’ from ‘forest’), some separating homophones (e.g., the grave accent in ‘où’ has no effect in pronunciation, and it only serves to distinguish the word from ‘ou’), and others have a clear effect on the sound quality (e.g., the acute accent in ‘é’ or the cedilla in ‘ç’).

The status of diacritical characters varies between languages, for example ‘é’ is considered a letter in its own right in Hungarian, but in French or Spanish, it is viewed as a modified ‘e’. But does the VWFA represent it as a separate character? In theory, this could be tested by letter similarity priming effects. When the replaced letter is sufficiently similar to the original one, such a modified prime could facilitate the processing of the intact target (e.g., for the target word *OBJECT*, the primes *object* and *obiect* work equally well, both outperforming *obaect*; Marcet & Perea, 2017, 2018). Recently, letter similarities were tested with isolated diacritical vowels and words containing diacritical letters in French (Chetail & Boursain, 2019). It was found that priming non-diacritical letters with diacritical ones had similar cost to replacing them with unrelated letters (RT pattern example for the target vowel *A*: *a* < *â* = *z*, and similarly for the target word *TAPER: taper* < *tâper* = *tuper*). Based on this, the authors argued that diacritical letters must have separate representations.

A similar pattern was also found in Spanish, when Perea et al. replicated the French experiment (RT pattern example for the target vowel *A*: *a* < *á* < *é*, and for the target word *FELIZ: feliz* < *féliz* ≤ *fáliz*), but with the diacritical vowel lying more in between the identity and the unrelated conditions (Perea, Fernández-López, et al., 2020). This pattern deviates from the French study in that the diacritical vowel does not hinder processing as much as the unrelated condition. This could be explained by the less important role of diacritics in Spanish, as they never alter the vowel sound. Crucially, Perea et al. also tested the other direction and revealed a different pattern when priming diacritical targets with non-diacritical primes, namely, that the omitted diacritic performed just as well as the identity prime (RT pattern example for the target vowel *Á*: *á* = *a* < *e*, and for the target word *FÁCIL: fácil* = *facil* < *fecil*). The lack of a processing cost when omitting diacritics could be again explained by the fact that Spanish makes no phonological distinction between diacritical and non-diacritical vowels. This explanation, however, is refuted by the fact that the same pattern was replicated in Finnish (RT pattern example for the target vowel word *PÖYTÄ: pöytä* = *poytä* < *paytä*; Perea, Hyönä, et al., 2022), a language that uses diacritics to distinguish between very distinctly sounding front and back vowels (e.g., o /o/ vs. ö /ø/).

The above detailed studies could be suggestive that only visual factors influence the perception of diacritics. This would be in line with computational models of reading, as they generally suggest, that addition of information differs from the absence of it. Features incompatible with a letter could cause direct inhibition (e.g., the acute accent serves evidence against the base letter a). This type of inhibition, however, was disproved by Rey et al. (2009)

in an ERP investigation. Another explanation could be based on Bayesian inference, where the presence of a feature is much stronger evidence, than the absence of it. This is implemented in the noisy-channel model of Norris and Kinoshita (2012), and tested on diacritical katakana characters (Kinoshita et al., 2021), stating that the diacritic is a purely visual feature, and its linguistic effects do not show up in the early stages of processing. In fact, they argue, that diacritical priming is akin to other letter similarity effects, where a single feature differentiates two letters, reproducing the same asymmetry (e.g., A primes Á but Á does not prime A the same way as F primes E but E does not prime F).

This theory is very straightforward, however there is some evidence against such a clear-cut account, namely that the size of the priming effect seems to depend on the linguistic function of the diacritic. The introduction of extra diacritics was tested in English, the readers of which are not expected to have internal representations of accented letters (Perea, Gomez, et al., 2022), and although the effect was present, its size was very small. So far the effect sizes are 7 ms in English (nórth – NORTH), 17 ms in Spanish (féliz – FELIZ; Perea, Fernández-López, et al., 2020), and 50 ms in French (néveu – NEVEU; Chetail & Boursain, 2019), suggesting that the linguistic function strongly modulates the similarity effect. It is also reasonable to assume that the linguistic modulation would depend on orthographical depth, as with shallow orthographies, the grapheme-phoneme conversion is independent of context. In this case, phonological abstraction is easily 'outsourceable' to the level of visual processing. This could be tested by investigating the similarity effect with varying linguistic roles of diacritics in the same language.

1.5 Masked repetition priming in reading research

For over three decades, one of the most used tools in word recognition research has been masked repetition priming (Grainger & Jacobs, 1999). It is based on the standard priming effect, i.e., after being presented a prime word, people tend to respond more quickly to following target words if they are semantically related to the prime (DOCTOR – NURSE), and more slowly if they are unrelated (TREE – NURSE). The problem with this setup is that the overt nature of priming can influence the participants' strategy and alter the validity of the results. Masked priming on the other hand resolves this issue by presenting the prime for a very brief period of time (usually 50 ms), and further reducing its saliency by incorporating it between a forward and/or a backward mask (e.g., a row of # marks). The prime is thus rendered subliminal, and its effects on the following target stimulus are automatic and not influenced by participant strategy (Forster et al., 1987). This effect is understood to be mostly prelexical – only affecting orthographic processing steps –, although the presence of a backward mask can increase the prime to target interval and allow for lexical and semantic effects too. Crucially, the short presentation means that the nervous system cannot repeatedly sample the visual information and is bound to work with the information present upon initial perception. The theory behind priming is that the prime can, to some extent, preactivate the reading network, leading to increased performance (faster RT, increased accuracy) when the target is the same word as the prime, and could cause interference if the words are different (Dehaene et al., 2001; Holcomb & Grainger, 2007). If we see stimulation, it means, that the prime was processed, and upon initial processing, it is represented similarly to the target. The size of the effect thus reflects the processability and the similarity of the prime.

2 Aims

No studies so far have presented a detailed view on the psychophysical effects of word orientation on automatic processing. We aimed to conduct a repetition priming experiment that describes the relationship between word orientation and readability, by rotating only the prime stimuli, and presenting normally oriented target stimuli. Our goal was to utilize rotations all around the circle (as opposed to previous works focusing only on a smaller range and one direction), to avoid overgeneralization of rotation effects. Since correcting cognitive mechanisms are expected when presented with rotated stimuli (e.g., mental rotation), we opted to modulate the duration of prime presentations, and thereby also modulate the facility of such correcting mechanisms. Based on the theory and research of Dehaene and Cohen (Cohen et al., 2008; Dehaene et al., 2005), we expected to find a steep change in the presence of rotated priming effects; their prediction is that the effect should essentially turn off above 45° rotation, and this pattern would not be affected by priming duration. We also aimed to assess whether any rotated word priming effects can be explained solely by the presence of the right letters, without the correct orthographic information, by reversing the letter order of the prime stimuli in a second experiment.

Another yet understudied area of visual word form processing has been the theory behind diacritical letters. Although they are used widely, most mainstream research focuses on English writing, and fails to capture the mental representation of modified characters. Since many languages use diacritical letters for various linguistic reasons, it is necessary to study multiple of them to gain a unified view of diacritic processing. To supplement the present literature (Chetail & Boursain, 2019; Kinoshita et al., 2021; Perea, Fernández-López, et al., 2020; Perea, Gomez, et al., 2022), we intended to study the acute accent in Hungarian, a language with a unique combination of ubiquitously used diacritics and very shallow orthography. By varying the presence or absence of diacritics on different base letters in a masked priming paradigm, we could target separate linguistic functions of the same diacritic. The priming pattern could either show asymmetry as evidence for the superiority of visual factors, or it could be more symmetrical, highlighting the importance of the linguistic roles. Since ‘o’ and ‘ó’ are closer phonetically, than ‘a’ and ‘á’ (see rationale in the methods section, and also Figure 4), we expected that if present, the linguistic pattern would be more pronounced in the latter case.

3 Materials and methods

3.1 Rotated word priming (Benyhe & Csibri, 2021)

The design was based on the first experiment of Harris et al. (2008), but instead of images of objects, our stimuli were written words. In this paradigm, the masked priming technique was utilized with backward masking and responses consisting of reading the target word aloud. The primes were either the same or unrelated to the target word and were rotated in various degrees. In Experiment 1, we tested a range of different priming durations to modulate the orientational effect, and in Experiment 2 we used reversed letter order for the prime words to test for the separation of letter and word orientations in word form processing (Figure 3).

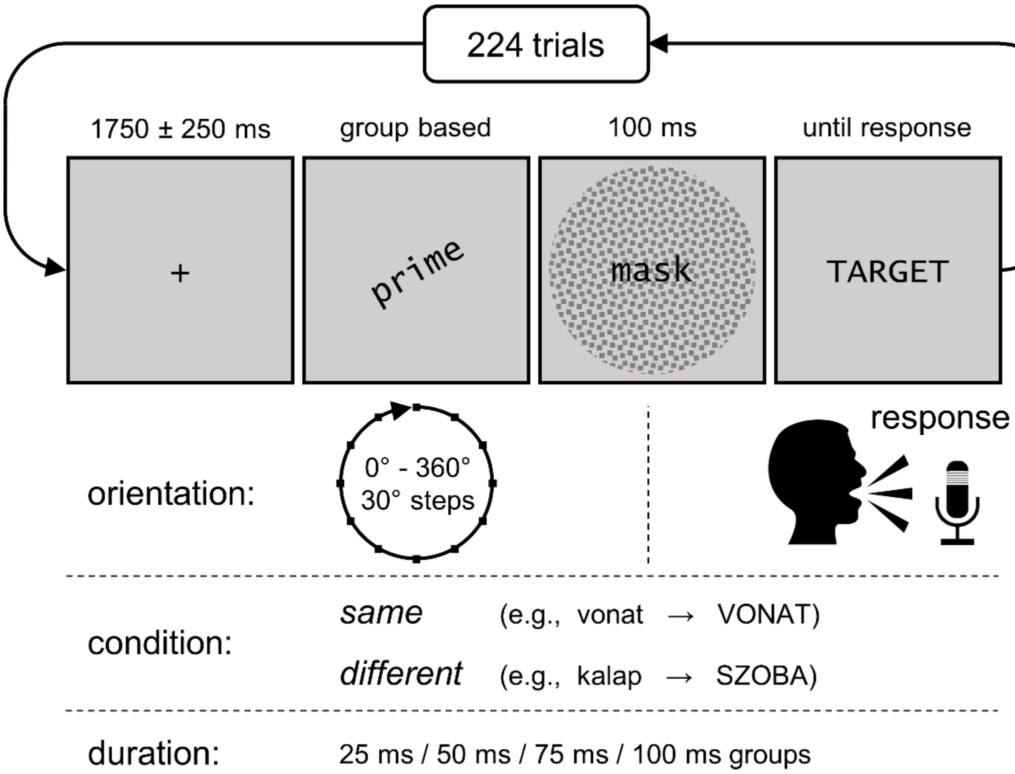


Figure 3 Design of Experiment 1 of the rotated word paradigm. Each trial starts with a fixation cross of variable interval, followed by the prime stimulus. The prime appears in one of twelve possible orientations and can be of two conditions: a *same* prime is followed by an identical target word, whereas a *different* prime is followed by an unrelated target word. The prime duration varies between groups. The prime is followed by a circular mask for 100 ms, after which the target stimulus appears, staying on screen until verbal response. Responses are validated offline for accuracy and response time. The design of Experiment 2 is the same as the 50 ms group of Experiment 1, except for the reversed letter order of primes (e.g., a *same* trial could be tanov→VONAT).

3.1.1 Experiment 1

Participants

A total of 53 students (34 female, 19 male, mean age = 25.6 years) from the University of Szeged were recruited to participate. All of them were native Hungarian speakers, with no declared reading or speech disabilities and normal or corrected-to-normal vision. The experiment was approved by the Regional Research Ethics Committee of the University of Szeged (267/2017-SZTE) and was conducted in accordance with the Declaration of Helsinki. The participants were randomly split into four groups of different priming durations (there were 10, 17, 16, and 10 participants in the four groups, respectively). Simulation based on pilot data showed that a participant count of 10 in the current design would yield 80% power for an effect size of 30 ms.

Stimuli

Stimulus words were selected from the Hungarian National Corpus (HNC, Oravecz et al., 2014) as the 336 most frequent Hungarian noun lemmas with a length of 5 or 6 characters (frequency range: 196-17/million, Zipf-frequency range: 4.23-5.29). To create the stimuli, the words were drawn in black color with a monospaced typeface (`Lucida Console`) at the center of a light gray background. The same words were used for both prime and target stimuli. To minimize visual similarity, primes were drawn with lowercase letters, and targets with uppercase letters.

Masks had to be designed in a way to conceal prime orientation, as the standard row of hashmarks would not suffice. In order not to give off any hints on prime orientation, masks were produced by uniformly scattering 300 Hungarian characters with random orientations on the center of the screen, cropped to a circle. To reduce any effect of familiarization to the masks, five were generated and cycled throughout a session, and they were drawn in random orientations.

Procedure

The experiment was designed and run in MATLAB with the Psychtoolbox 3 extension (Kleiner et al., 2007), under Microsoft Windows. Trial lists were created in three versions, by permuting thirds of the complete stimulus set in a Latin square manner (each third would switch roles as primes of the *different* condition, targets of the *different* condition and primes and targets for the *same* condition). This way each word would only be presented in one trial during one session, and each role would be roughly balanced between unique words after multiple sessions, as lists would be assigned randomly at the start of each run. The order of the trials was permuted randomly, and the two priming and twelve orientation conditions were balanced evenly. One run consisted of 224 trials (plus five practice trials) and lasted on average 10 minutes. Trial structure is summarized in Figure 3.

The experiment was carried out in a darkened, quiet room to reduce the possibility of distractions. After giving written consent, participants were seated at 57 cm distance from an Asus PG248Q 120 Hz monitor used for stimulus presentation and had a Rode NT-USB microphone in front of them to capture verbal responses. Participants were instructed to read target words aloud as quickly and accurately as possible. The computer recorded response onset and paced the experiment automatically. The recorded utterances were validated offline for accuracy and precise RT measurement.

Data analysis

Trials with inaccurate utterances or RT above 1000 ms were excluded from the analysis. Analyses were performed in R (Version 3.6.2) using the lme4, and emmeans packages. Separate generalized linear mixed-effects models (GLMM) with inverse Gaussian function and log link were fitted to the RT data from each experimental group (Bates et al., 2015; Lenth et al., 2018; R Core Team, 2021). The following model structure was used to fit the models (expressed in Wilkinson notation):

```
RT ~ condition * orientation + log(n_trial) + (1 + condition
     | participant) + (1 | item)
```

where both condition and orientation are fixed factors and (the natural logarithm of) trial number is continuous variable.

Fixed effects were evaluated with Wald tests, and the priming effect for all orientations was calculated as Tukey-corrected estimated marginal mean (EMM) contrasts between the *same* and *different* conditions. Plots were created with the ggplot2 package (Wickham, 2016).

3.1.2 Experiment 2

Participants

Fourteen students (9 female, 5 male, mean age = 21.7 years) from the University of Szeged were recruited for the second experiment. All were native Hungarian speakers, with no declared reading or speech disabilities and normal or corrected-to-normal vision. The experiment ran under the same ethical license as Experiment 1.

Stimuli

Stimuli and masks were created similarly to Experiment 1, with the exception, that all prime words were drawn with reversed letter order.

Procedure and data analysis

All procedures and analysis steps were performed as in Experiment 1.

3.2 Diacritic priming (Benyhe et al., 2023)

Two experiments were designed to evaluate the effects of the acute diacritic used on two sets of vowels to produce two distinct linguistic functions. Hungarian has a highly transparent orthography and a ubiquitous use of diacritics, with roles well defined by phonology (Figure 4). Each vowel has short and long versions, the latter of which is signaled by acute (or double acute) accents. The long and short versions are usually remarkably similar in their sound qualities and differ only in length (e.g., 'o' /o/ and 'ó' /oː/) but there are exceptions, when there is a contrast in quality as well (e.g., 'a' /ɒ/ and 'á' /aː/). In Experiment 1, we modulated diacritic presence on the letters 'o/ó' and 'u/ú' to change the length of the vowel; in Experiment 2, the modulation of the same diacritic produced an extra effect in vowel quality besides changing vowel length on the letters 'a/á' and 'e/é' (Figure 5).

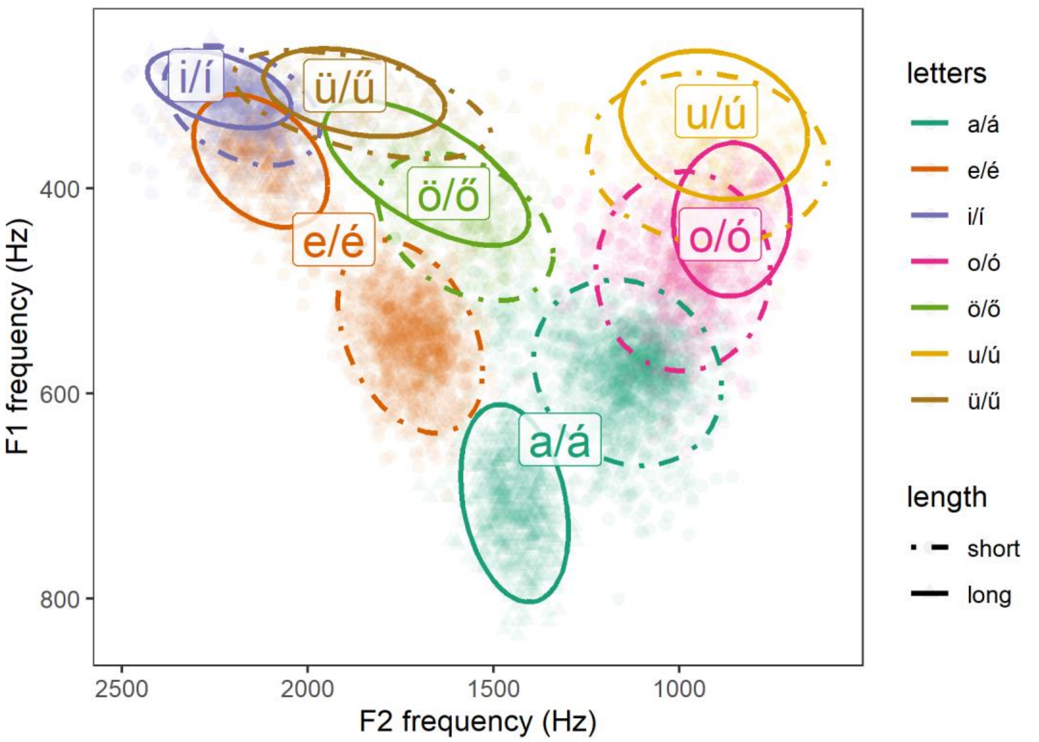
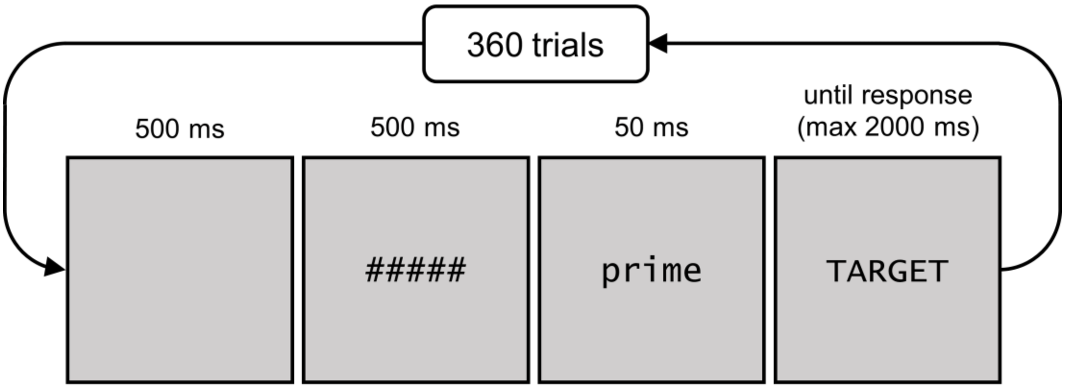


Figure 4 Formant distributions of the vowels of the Hungarian alphabet. Data from 6256 individual sounds of a male speaker are redrawn with permission from Abari et al. (2011). The first and second formant frequencies (F1 and F2) characterize the quality of an individual vowel sound, and ellipses are drawn at the 95% level for each letter. Related letters are plotted with the same color and their lengths are differentiated by line style. Whereas most short-long pairs have overlapping distributions, the letters of the 'a/á' and 'e/é' pairs are more distant from each other, indicating a contrast not only in sound length, but in sound quality as well.



examples of word stimuli		prime condition			target
		identity	***related***	***unrelated***	
target diacritic condition	***non-diacritical***	kapu	kapú	kapá	KAPU
		mese	mése	máse	MESE
	diacritical	róka	roka	reka	RÓKA
		néma	nema	nama	NÉMA

Figure 5 Design of both experiments of the diacritical paradigm. The task is lexical decision, so targets can be words or pseudowords. The trials can be of two main conditions based on the diacritical status of the key letter in the target stimulus: *diacritical* or *non-diacritical*. In Experiment 1, the key letters were ‘o’ or ‘u’ in the non-diacritical, and ‘ó’ or ‘ú’ in the diacritical condition, whereas in Experiment 2 these were ‘a’ or ‘e’, and ‘á’ or ‘é’ for the two conditions, respectively. The trials start with a 500 ms blank interval, followed by a 500 ms forward mask. The mask contains the same number of hashmarks (#) as the following prime and target. The prime stimulus is presented for 50 ms and can be one of three conditions: the *identity* prime is the same string as the target, the *related* prime has the diacritical status of its key letter flipped, and the *unrelated* prime has a completely different key letter with a flipped diacritical status. The prime is followed by the target stimulus that stays on until button press response or a maximum of 2000 ms.

3.2.1 Experiment 1

Participants

A total of 72 young adults (41 female, 29 male, 2 rather not say, mean age = 24.7 years) were recruited via the online platform Prolific Academia. All of them were native Hungarian speakers, with no declared reading or writing disabilities and normal or corrected-to-normal vision. The experiment was approved by the Ethics Committee of the University of Valencia and was carried out in accordance with the Declaration of Helsinki. The participant count was determined based on the recommendations of Brysbaert and Stevens (2018).

Stimuli

One hundred and eighty target words were selected from the HNC, filtered for two-syllable lemmas, 4 to 6 characters in length. Each word contained at least one of the key letters, ‘o’ or ‘u’ in the non-diacritical, and ‘ó’ or ‘ú’ in the diacritical target condition, with key letter occurrences balanced in the list (45 instances each). Importantly, the key letter never occurred in the first position, and the other vowel in the word was never diacritical. Words were selected in a way to minimize the differences in word frequency and key letter position between the diacritical and non-diacritical conditions (mean Zipf-frequency: 3.96 and 4.49, respectively; mean key letter position: 3.5 and 3.1, respectively).

Prime words were derived from target words. Each target had three associated prime words one for each of the three priming conditions: identity primes were the same as target words; related primes were different from target words only in the diacritical status of the key letter (e.g., ‘o’ becomes ‘ó’ or ‘ú’ becomes ‘u’); and unrelated primes had a completely different vowel with opposite diacritical status replacing the key letter (e.g., ‘o’ becomes ‘á’ or ‘ú’ becomes ‘o’). The primes were created in a way to produce pseudowords in the related and unrelated conditions, and targets for which this was not possible were excluded.

For the lexical decision task, an equal number of pseudoword targets had to be created. These were produced by generating pseudowords around the vowel cores of real words, based on letter pair and triplet probabilities observed in the HNC. The associated primes were generated similarly to those of word targets.

Each stimulus was created by drawing the words in black color and a monospaced typeface (Lucida Console) on the middle of a light gray background. Targets were always presented in uppercase, and primes in lowercase. See scripts, documentation and complete stimulus lists in the data repository associated with Benyhe et al. (2023).

Procedure

The experiment was coded in PsychoPy3 (Peirce et al., 2022) and ran online via the online server Pavlovia. To balance out the priming conditions among individual items, three stimulus lists were created with a Latin square design, and each was run on equal number of participants. Trial order was permuted randomly, and target diacritic and priming conditions were balanced evenly throughout the experiment. After 20 practice trials (aided with accuracy feedback), the experiment consisted of 360 trials (180 words and 180 pseudowords, see trial structure in Figure 5), with self-paced breaks after every 60 trials. One run lasted approximately 20 minutes.

The recruited participants gave informed consent, and ran the experiment on their own computers (for validity of online experiments, see Angele et al., 2022). Participants were instructed to complete the experiment in a quiet environment with no distractions. The task was to decide if the appearing string is a real Hungarian word or a nonsense pseudoword, as quickly and accurately as possible, by pressing keys on the keyboard (‘X’ for pseudoword, ‘M’ for real word). Response accuracy and RT were recorded.

Data analysis

Separate analyses of RT and accuracy data were performed in R with the brms package for Bayesian linear mixed-effects modeling (Bürkner, 2017; R Core Team, 2021). Trials with shorter than 250 ms RTs were excluded from the dataset, and only correct responses were used for the RT analysis. Bayesian mixed-effects models (5000 iterations) were fitted with the following structure:

```
dependent_variable ~ diacritic * prime + (1 + diacritic *
    prime | participant) + (1 + diacritic * prime | item)
```

where diacritic and priming conditions were coded as fixed factors (with the non-diacritical similar condition as reference). The RTs were modeled with Gaussian distribution of inverse transformed data (-1000/RT). Accuracy was modeled with Bernoulli distribution. Effects with a 95% credible interval (CrI) not crossing zero were regarded as significant. Plots were made with the ggplot2 package (Wickham, 2016).

3.2.2 Experiment 2

Participants

A second set of 72 young adults (36 female, 34 male, 2 rather not say, mean age = 24.6 years) were recruited via Prolific Academia. All of them were native Hungarian speakers, with no declared reading or writing disabilities and normal or corrected-to-normal vision. The experiment ran under the same ethical license as Experiment 1.

Stimuli

Stimulus lists were created in a similar fashion to Experiment 1. Crucially, the key letters were changed to ‘a’ and ‘e’ in the non-diacritical, and ‘á’ and ‘é’ in the diacritical target conditions. The 180 words were again selected in a manner to minimize differences in frequency and the position of the key letter between the diacritical and non-diacritical conditions (mean Zipf-frequency: 4.75 and 4.87, respectively; mean key letter position: 3.2 in both conditions).

Procedure and data analysis

Procedures and analyses were performed in the same way as for Experiment 1.

4 Results

4.1 Rotated word priming (Benyhe & Csibri, 2021)

4.1.1 Experiment 1

Out of the 11872 observations, 521 (~4%) were excluded due to inaccurate response, failure to respond within the time window, or recording error. Separate GLMMs were fitted for each priming duration group with successful convergence. The Wald tests confirmed a significant interaction between prime orientation and priming condition in each group, justifying the pairwise comparisons along the orientations ($W_{25ms}(11) = 44.1$, $p < .001$; $W_{50ms}(11) = 260.3$, $p < .001$; $W_{75ms}(11) = 164.6$, $p < .001$; $W_{100ms}(11) = 77.2$, $p < .001$). EMMs and priming effects are summarized in Table 1, and mean RTs are plotted in Figure 6.

Table 1 Estimated marginal means of RT along conditions and orientations in ms. For each experimental group, priming effect is calculated as the contrast between *same* and *different* conditions. Asterisks mark the Tukey-corrected significance of priming effects: * p < .05, ** p < .01, *** p < .001.

			Prime orientation											
			0°	**30°**	**60°**	**90°**	**120°**	**150°**	**180°**	**-150°**	**-120°**	**-90°**	**-60°**	**-30°**
Experiment 1	**25 ms**	*same*	470	495	486	484	493	487	497	493	497	503	486	473
		different	500	491	492	486	487	492	486	489	484	491	482	493
		priming	29***	−4	7	2	−6	5	−11	−4	−12	−12	−3	20*
	50 ms	*same*	470	480	486	507	515	527	521	514	517	510	484	469
		different	555	549	538	515	518	520	522	532	519	537	537	546
		priming	85***	69***	51***	8	4	−7	1	18	1	27**	53***	76***
	75 ms	*same*	457	481	488	501	504	505	502	495	503	489	466	458
		different	540	537	528	521	518	516	519	525	514	510	524	538
		priming	83***	56***	39***	20*	13	11	17*	30**	10	21*	58***	80***
	100 ms	*same*	471	474	465	497	503	504	494	500	498	483	473	475
		different	534	538	531	520	519	510	520	515	514	512	548	536
		priming	63***	64***	66***	23*	16	6	25*	15	15	29*	74***	60***
Exp 2	**reversed**	*same*	512	495	499	494	493	493	492	495	493	488	497	500
		different	495	503	493	488	491	489	492	496	493	487	494	493
		priming	−17**	8	−6	−5	−2	−4	0	1	0	−1	−2	−7

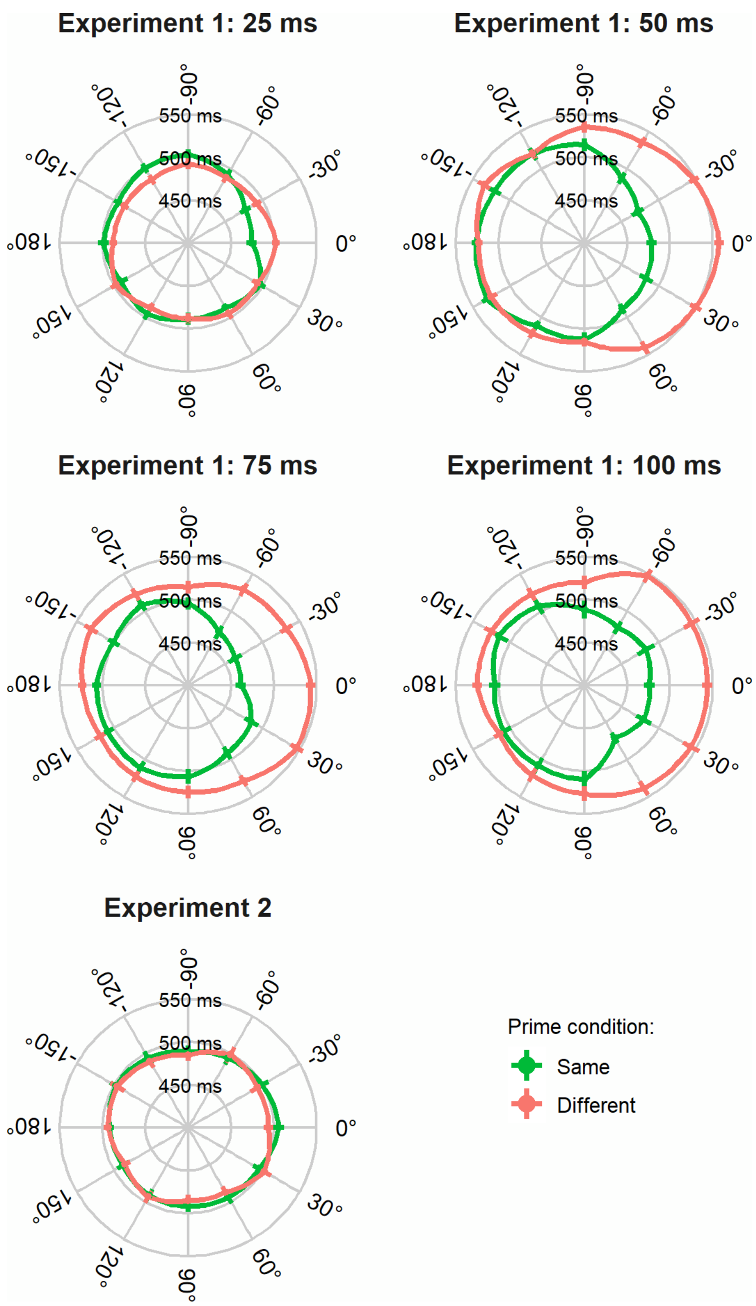


Figure 6 Mean response times ± 1 S.E.M. as a function of prime orientation, colored by priming condition. The orientation labels are rotated in the same degree as the denoted prime rotations.

Positive priming effects were found in all priming duration groups. Whereas this effect is only present around normal orientation (0° and -30°) in the 25 ms group, the range spreads and stabilizes in the longer priming groups (sizable effect from -60° to 60°) and can still be detected in the perpendicular orientations (±90°). There are also incidental cases of statistically significant priming effects in far oriented or upside-down orientations (e.g., a 30 ms effect at -150° in the 75 ms group).

4.1.2 Experiment 2

The GLMM was fitted to 3037 observations, 99 trials (~3% of all) were rejected. With the reversed prime words, there was no significant main effect of prime condition, and there were no apparent positive priming effects in any orientation (Table 1, Figure 6). The multiple comparisons of EMMs indicated a small negative priming effect at 0°, although this effect is questionable due to the lack of significant main effect and interaction in the Wald tests.

4.2 Diacritic priming (Benyhe et al., 2023)

4.2.1 Experiment 1

After deleting trials with too short responses, and eliminating the data from one word item, which performed under 50% accuracy, the remaining 12864 trials were used for the accuracy model, out of which the 12089 correct responses were used for the RT model. Mean RT and accuracy data are summarized in Table 2.

Table 2 Mean response times in ms and mean response accuracy in parentheses across each condition. Note that only the data from word trials were used in the analyses, and pseudoword data are only shown for consistency.

		Word targets		**Pseudoword targets**	
		Non-diacritical	**Diacritical**	**Non-diacritical**	**Diacritical**
Experiment 1	*identity*	632 (.947)	665 (.926)	704 (.963)	727 (.958)
	similar	643 (.952)	661 (.923)	707 (.954)	732 (.957)
	dissimilar	656 (.948)	681 (.932)	712 (.965)	734 (.953)
Experiment 2	*identity*	638 (.971)	640 (.964)	709 (.967)	731 (.956)
	similar	650 (.949)	641 (.970)	718 (.975)	743 (.954)
	dissimilar	666 (.944)	652 (.955)	719 (.978)	733 (.964)

The RT model (Figure 7) indicated a significant advantage of the *identity* condition (over *similar*, b = -0.03, SE = 0.01, CrI [-0.05, -0.01]), modulated by diacritical status (interaction, b = 0.03, SE = 0.01, CrI [0.00, 0.05]), thus the identity advantage was only present for non-diacritical targets. Conversely, the *dissimilar* condition had a disadvantage (compared to *similar*, b = 0.03, SE = 0.01, CrI [0.01, 0.05]), but this was unaffected by diacritics (interaction, b = 0.00, SE = 0.01, CrI [-0.02, 0.03]) showing a similar pattern for both diacritical and non-diacritical targets.

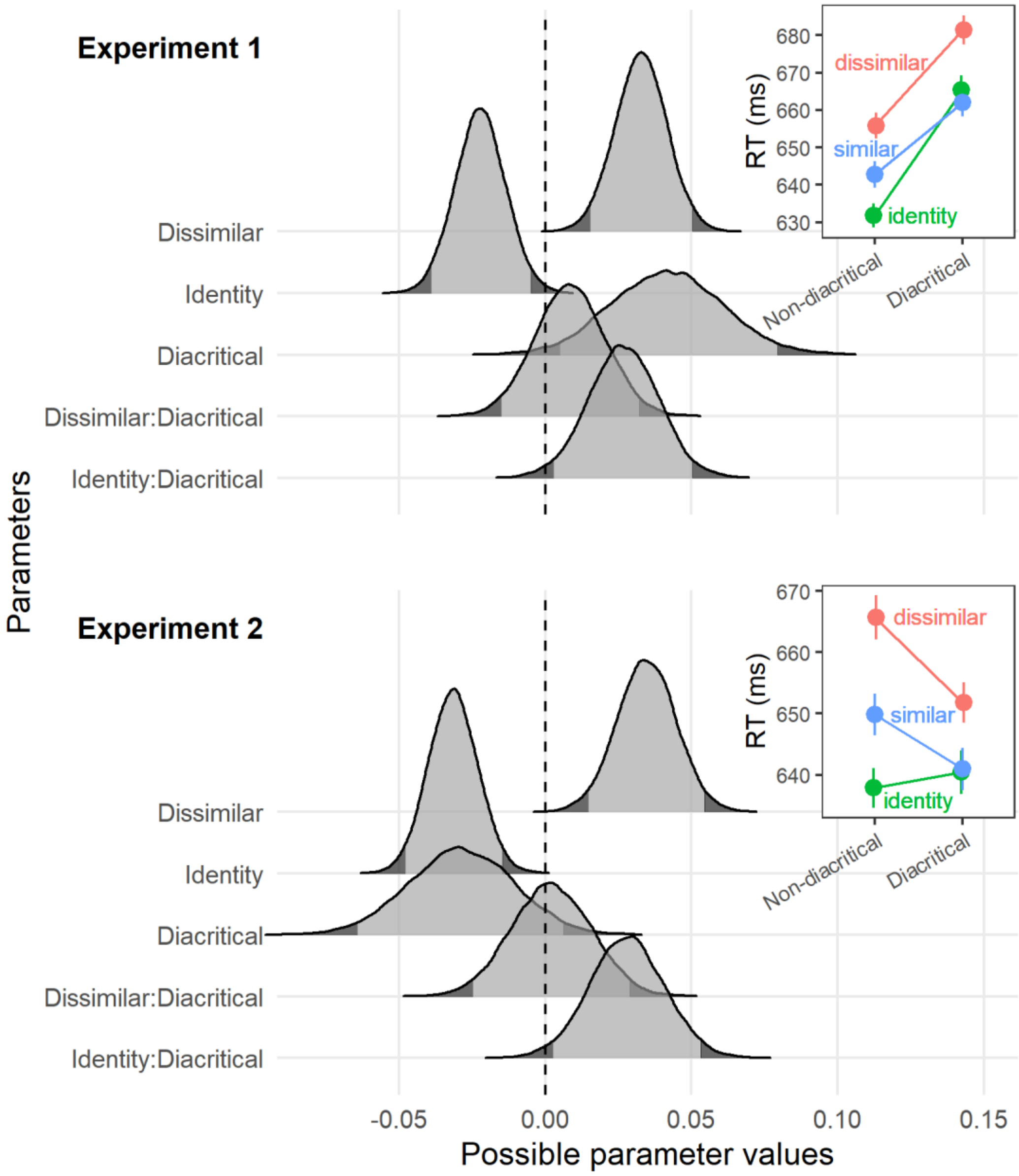


Figure 7 Posterior distribution of parameter values in the Bayesian RT models. Parameters with their 95% credible interval (light gray shading between the tails) not crossing zero are regarded as significant. The insets show mean response times ± 1 S.E.M., colored by priming condition.

The accuracy model (Figure 8) found no significant effect of the *identity* or *dissimilar* conditions (b = -0.10, SE = 0.16, CrI [-0.41, 0.21] and b = 0.17, SE = 0.22, CrI [-0.26, 0.59], respectively), and their interaction with diacritical status was also insignificant (b = -0.09, SE = 0.21, CrI [-0.31, 0.49] and b = 0.17, SE = 0.22, CrI [-0.26, 0.59]).

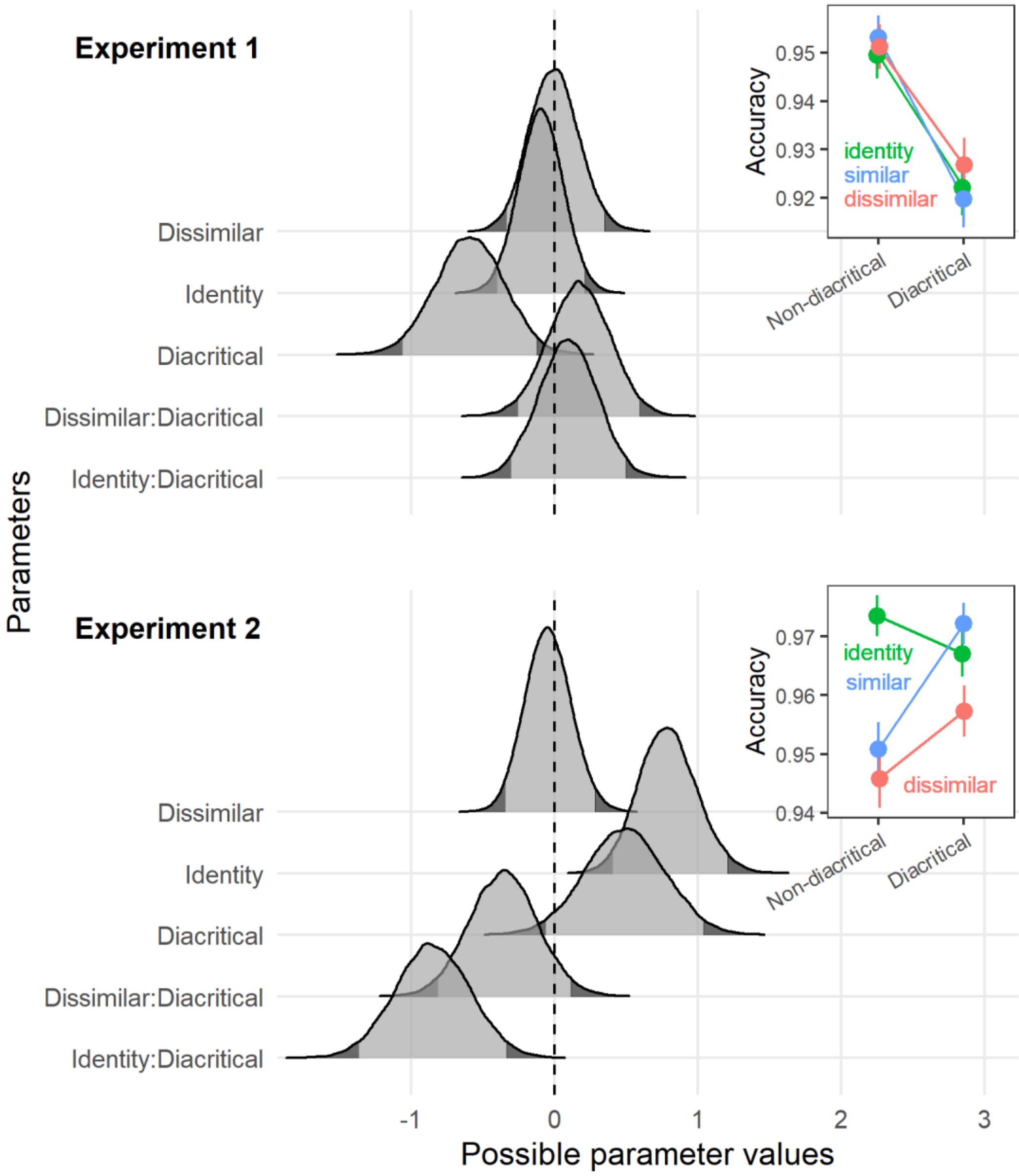


Figure 8 Posterior distribution of parameter values in the Bayesian accuracy models. Parameters with their 95% credible interval (light gray shading between the tails) not crossing zero are regarded as significant. The insets show mean accuracy ± 1 S.E.M., colored by priming condition.

4.2.2 Experiment 2

Trials with responses below 250 ms were eliminated, and the remaining 12927 trials were used for fitting the accuracy model, out of which the 12424 correct responses were used for the RT model. Accuracy and RT means are summarized in Table 2.

Analogous to the previous experiment, the *identity* condition was once again faster than the *similar* condition (Figure 7; b = -0.02, SE = 0.01, CrI [-0.04, -0.01]), but only for the non-diacritical targets (interaction, b = 0.03, SE = 0.01, CrI [0.00, 0.05]). A disadvantage of *dissimilar* primes was present (b = 0.03, SE = 0.01, CrI [0.02, 0.05]), regardless of diacritical status (interaction, b = 0.01, SE = 0.01, CrI [-0.02, 0.03]).

The *identity* advantage was also present in the accuracy model (Figure 8; b = 0.79, SE = 0.20, CrI [0.40, 1.20]), with the same interaction pattern (b = -0.85, SE = 0.26, 95% [-1.37, -0.34]). The disadvantage of the *dissimilar* condition was, however, absent from the accuracy model (main effect: b = -0.05, SE = 0.16, CrI [-0.35, 0.28]; interaction: b = -0.36, SE = 0.24, CrI [-0.83, 0.11]).

5 Discussion

5.1 Rotated word priming (Benyhe & Csibri, 2021)

We conducted two masked repetition priming experiments with rotated words as primes to test the effects of orientation on orthographic processing. In the first experiment, we found that the priming effect was strongly dependent on orientation, which is in contrast with the findings of Harris et al. (2008) regarding object recognition in a similar priming paradigm. This highlights the divergence of orthographic processing from object recognition and shows a reduced tolerance towards transformations in word form processing. Indeed, during learning to read, the brain must 'unlearn' the property of mirror equivalence, that is otherwise present in object recognition (Dehaene et al., 2005), so it can be assumed that the two subsystems can perform differently in case of other transformations too. Also, words are represented in the brain as a directed sequence of letters (Grainger, 2018), so the ability to read a word should be dependent on finding the main axis, along which the code exists.

Curiously, our results showed that with the shortest duration, the priming effect was only significant at 0° and 30° (~20 ms), but with longer durations it opened up to ±60° (~60 ms) and was even detectable to a lesser extent at ±90° (~20 ms). Although priming effects are expected to grow approximately linearly with prime duration (Grainger & Jacobs, 1999), this was not the case in our study; neither for normally oriented primes, nor in any other direction. Thus, it seems that there is a time-dependent component of rotation-tolerance which quickly reaches a plateau (from 50 ms on). The range of the rotated priming effect is roughly in line with the LCD model's predictions (breakdown above ~45°; Cohen et al. 2008) but is a bit wider (at least 60°) and more gradual in its decay.

When trying to account for rotation effects in the SERIOL model's framework, Whitney (2002) assumed, that the letter nodes received weaker input due to mental rotation. This would break the temporal code which could not be passed to the next level of processing within a single oscillatory cycle. This reasoning however ignores the finding, that single letter recognition is remarkably resilient to rotations (Koriat & Norman, 1989; Perea et al., 2020). Also, according to the model, location gradient is built from relative visual acuity within the visual field, an idea strongly coupled with the anatomical properties of the retina that is seemingly incompatible with mental rotation (the observer cannot imagine how acuity would change if the image of a letter projected on another part of the retina). Even if rotated words need more oscillatory cycles or sub-cycles (the temporal grid in the SERIOL model, 200 ms and 25 ms, respectively, proposed by Whitney, 2002) to properly represent the orthographic code, we should see a more stepwise pattern: all orientations that are processable in the same number of cycles should behave similarly. In our results, we could say that the first step is present, when transitioning from 25 ms primes to the 50 ms duration, but then it would follow that a further increase yields a wider rotation tolerance. Furthermore, the shortness of the prime-target interval, and the masking renders conscious efforts at stimulus normalization improbable (i.e., mental rotation proposed by Whitney et al. [2002] and agreed on by Cohen et al. [2008]). Also, the mixed nature of our design (as opposed to block designs) further reduces the possibility of participants anticipating the orientation. In our opinion, the orientational priming pattern reflects the amount of information readily and automatically available in the early stages of processing, rather than a cognitive effort to mentally rotate the word form back to its canonical orientation. Thus, the SERIOL model's explanation cannot fully account for the effects seen with masked priming.

Other computational models, such as the overlap model, do not explicitly state their predictions with rotated word forms (Gomez et al., 2008). If we try to apply the overlap model, we should only take the letters' horizontal position into account, or we should expand the position representations in the second dimension. Either way, we have to assume, that the overlaps between letter representations increase to an extreme at ±90° orientation, at which point, the normal and reversed order primes would have equal representations. In the reversed prime experiment, however, we did not see any sign of priming in the perpendicular orientations.

An alternative explanation of the rotation effect would be that it is caused by the effort to find the axis along which to extract the code. This could take the form of opening an attentional window, defining the boundaries of the word and then analyzing its components. In case of normally oriented primes, for which the visual system is prepared, this needs no extra effort, hence we can see a priming effect with the 25 ms primes at 0° orientation but not much further. Longer prime presentations, however, could possibly allow for a feedback cycle with updated expectations of orientation. An interpretation of this is the update of the frame of reference in which information is coded. This is possible through the mental capacity termed perceptual upright (Dyde et al., 2006), defining the reference frame in which the ambiguous rotated character 'p/d' is recognized as either a *p* or a *d*. The perceptual upright was found to be strongly influenced by the orientation of the visual background (having approximately half the weight of the body axis, the purely egocentric reference frame). In our experiments, the visual ‘background’ is the context in which the letters are perceived, i.e., the orientation of the whole word. Note that in our case the orientation of the word and the constituent letters is the same, unlike in recent work focusing on the rotation tolerance of letter detectors (Fernández-López et al., 2021, 2022). Such divergence of letter and whole word axes could account for the reduced tolerance (30°) found with parafoveal preview and masked priming experiments.

The proposed linked nature of word and letter orientation also explains why there was no priming at 180° in the second experiment. In the *same* condition, although individual letters are all rotated 180°, in each position, the letter identities are shared between the prime and the target. As letter recognition withstands such extreme rotations (Koriat & Norman, 1989), we would expect the priming pattern to flip in Experiment 2, if the letter and word orientations are coded independently. As this was not the case, we conclude that the resilience of orthographic processing towards rotations is greatest when letter and word orientations agree.

Finally, we should explain why there is an extreme cost for unprimed lexical decision (Cohen et al., 2008; Koriat & Norman, 1985) but a more robust priming effect for rotated words up to 90°. We find that the best approach here would be Bayesian. The initial evidence in a rotated prime word can be decoded to some extent and enriches the prior probabilities before sampling the target word as usual, hence the priming effect. In case of an unprimed paradigm, where the target word itself is rotated, we have to account for the rate at which evidence accumulates. We argue that this could be affected strongly by rotation: the initial guess forced by priming is correct, but the confidence for naming or lexical decision builds up more slowly. Therefore, even a 50 ms long rotated prime can enhance target processing, whereas the same rotation would inflict delays longer than 50 ms with rotated targets.

Despite our efforts, some intriguing results remain unexplained. Firstly, the appearance of seemingly robust priming in far orientations of Experiment 1 is incompatible with our current understanding of the VWFA’s capabilities. For now, as it does not fit into the overall pattern, we must rule it out as a possible sampling error. Furthermore, the inverse priming effect in the normal orientation of Experiment 1 is puzzling, as we do not expect greater interference with reversed letter order than with a completely unrelated set of letters (see the reversed anagram effect with repetition blindness, Morris & Still, 2012). In future experiments, we will have to see if these findings hold up and should be taken into account when designing models of orthographic processing.

5.2 Diacritic priming (Benyhe et al., 2023)

We conducted two masked priming experiments to test if the effect of added or missing diacritics is dependent on the linguistic function. In the first experiment, the diacritic only increased vowel length, whereas in the second, the diacritic also changed vowel quality. The two experiments yielded surprisingly similar results. While the unrelated condition always performed worse than the two other conditions, the identity advantage (decreased RT compared to visually similar primes) was only present for non-diacritical targets, regardless of the experiment. For example, the diacritical target word *RÓKA* has the following RT pattern: *róka* ≈ *roka* < *reka*; meanwhile the non-diacritical target *MOZI* is primed in the following pattern: *mozi* < *mózi* < *múzi*. Interestingly, the accuracy results did not parallel those of the RT: the cost of unrelated primes was completely absent, and the asymmetrical identity advantage was only present in Experiment 2. This could be interpreted as a ceiling effect, and in the following, we will only detail the effects in RT.

These findings suggest that the early processing and encoding of diacritics is independent of phonological features. The interference caused by *similar* primes is only present when the amount of visual information is greater in the prime, than in the actual target (ó⇏O), but not in reverse (o⇒Ó). Furthermore, this is not only true for deep orthographies but also for Hungarian which is orthographically remarkably shallow. Based on the triangle and generally the dual-route models (Coltheart et al., 2001; Seidenberg & McClelland, 1989), phonological effects are expected to be augmented in case of high orthographic transparency. This was not the case, and the fact that the changes in diacritical status produced the same priming pattern in Hungarian as in Spanish (Perea, Fernández-López, et al., 2020) or even English (Perea, Gomez, et al., 2022) points out that these effects arise along the visual route. This is in line with the predictions of the noisy channel model (Norris & Kinoshita, 2012), in that the visual system expects to receive noisy information, and is more prepared to fill in missing details than to ignore present ones. The presence of information serves greater evidence towards a specific letter identity, than the absence of information has against it.

We conclude that these findings support the idea that diacritical vowels are represented as separate letters. When the abstract letter identity has multiple forms (e.g., in the case of upper- and lowercase or italics), the same argument can be made that one version has more information than the other. For example, the uppercase letter *B* contains all features of the lowercase form *b*, but not the other way round. If these had separate representations, then one would expect to see asymmetrical priming effects (e.g., b⇒B but B⇏b). It was shown, however, that priming effects are case-independent (in other words symmetrical), both for letters with similar and dissimilar features in different cases (Kinoshita & Kaplan, 2008). Thus, if diacritical and base versions of the same vowel shared the same abstract letter representation, it would produce symmetrical effects. In contrast, we find asymmetry akin to that produced by visually similar but distinct letter identities (e.g., F⇒E but F⇏E; Kinoshita et al., 2021).

To account for the differences between languages, we propose that the diacritical prime to non-diacritical target interference depends on the development of the abstract letter representations. For languages where diacritics are in everyday use, the diacritical letters have a stable representation and produce a robust effect as in Hungarian (11-12 ms) or in Spanish (17 ms; Perea, Fernández-López, et al., 2020). In English, however the representation is expected to be much weaker, as the use of diacritics is less frequent, thus we find weak interference (7 ms). The results of Chetail and Boursain (2019) stand out from this comparison, as they found an extreme (50 ms) interference in case of French. This could be explained by the specific diacritics used in the experiments, as the Hungarian, Spanish and English studies used only the acute accent. The study regarding French, in contrast, employed visually more complex diacritics in approximately three quarters of the items, which could be more salient and produce more interference.

Importantly, these results only reflect the early stages of visual word form processing, and we should not rule out, that the diacritical letter detectors can give rise to differential effects based on the specific linguistic functions. Priming effects only confirm the presence of such letter detectors, but the way their activation weighs in the decoding process cannot be decided with this methodology and should instead be tested with unprimed paradigms. Perea et al. (2022) has already published promising results utilizing a semantic categorization task, where the omission of diacritics entailed a reading cost in German but not in Spanish. For more decisive results on the nature of diacritics, subsequent studies should compare languages or different sets of diacritics in such unprimed tasks.

5.3 General discussion

Reading research is an exciting and growing field of the cognitive sciences. Studies often focus on small sections of the complex task of reading (just like our contributions), but one should not lose sight of the bigger picture. The findings should eventually fit into a grand theory of reading, encompassing not only the visual, but the attentional, semantic and motor aspects too (as the OB1-reader model attempts to do, Snell et al., 2018). Crucially, such models should be compatible with the neuronal recycling account (Dehaene & Cohen, 2007), allowing the skill of reading to arise in a naïve neuronal environment. Computational models utilize predefined letter detectors, but the set of letter detectors in anyone's letterbox should depend on their personal experience, and their representation of written stimuli would be constantly shaped by statistical learning (Arciuli & Simpson, 2012; Pacton et al., 2001; Snell & Theeuwes, 2020). For example, readers of 18th century print would be well prepared to differentiate between the letters *ſ* (long s) and *f*, a hardſhip we modern readers can forget about. But when glancing into an old volume of Shakespeare, our visual system can more easily match the long s character with the abstract letter identity of *f*, as this form is basically missing from our experience with the letter *s*. One must assume that the brain region we now term VWFA has been there since ancient times, well before the practice of reading emerged, but was recruitable and available to meet the needs of this new task. And as we see, this new task keeps evolving.

Therefore, the capacities of the reading network will necessarily reflect the challenges it has to solve, for a Sumerian scribe and a modern reader alike. The fact that word reading has a substantial resistance to rotations just shows that the relative position between reader and text is not fixed. Some invariance towards text orientation is required to perform well in everyday situations, just as invariance is needed over letter case, style, position, etc. Similarly, a lot of the written content we read online is ripe with errors or often omits diacritics. The task our reading brain has to perform is to extract the intended linguistic information despite the typographical imperfections. What happened if instead, one would always see text in the same orientation, the same font, the same distance and the most immaculate spelling? In my opinion, the result would be overfitting, as seen in the simulation of Hannagan et al. (2021), when testing for rotations. The nervous system assigns its capacity to represent the variability observed in earlier experience and, as a result, is less equipped to find the nuance in unfamiliar items (e.g., see the other-race effect, Lindsay et al., 1991). This is to be expected in case of reading too and should explain most perceptual effects. This realization could inspire a new style of modeling, one that builds on the visual variability of prior experience about the written world.

6 Conclusion

In the first study, we successfully employed a novel scattered character mask to conceal the presentation of rotated prime words in a masked priming paradigm. Priming of word reading was detectable with prime rotations up to 90° but most effective below 60°. Importantly, this pattern was present for prime durations of 50, 75 and 100 ms, but not for 25 ms, where the priming was only present around the normal orientation. This moderate time-dependence shows that orientation resilience requires an extra step – possibly a feedback mechanism – but stays well within the bounds of automatic word form processing. This robust priming pattern questions earlier explanations involving mental rotation during reading rotated words. There was an absence of priming effect with reversed primes, showing that the rotated priming effect occurred after letter detection, and was truly orthographical. The reversed prime results also underline the importance of the alignment of word orientation, and the orientation of individual letters in word form processing.

In the second study, we modulated the presence and absence of diacritics of prime words in a masked priming paradigm with lexical decision. The experiments were done in Hungarian, an exceptionally transparent writing system, with native speakers. We found that similarity priming was not dependent on the linguistic function of the diacritic, despite the clear difference in pronunciation. The results revealed the same asymmetry as in other languages: priming occurs with non-diacritical prime and diacritical target, but not in the opposite direction. This finding confirmed that phonological factors do not have a role in masked priming, rather it is driven only by visual cues. Our results are in favor of Bayesian models of word form processing and agree with the theory that diacritical letters have detector units separate from the base letters.

Acknowledgements

Firstly, I would like to thank my co-supervisor Gyula Sáry, for inviting me to the department and giving me the opportunity and freedom with my work as a junior researcher and as a teacher. The physiology department has a unique sense of collegiality under his lead, which had a deep impression on me, and that I wish to carry on to my future workplaces.

I would also like to thank my supervisor, Péter Csibri, for guiding my first attempts in doing science. He always made me believe that all of it would get easy once I had my foot in the door and helped me immensely to get there.

I wish to express my utmost gratitude to Manuel Perea (Universitat de València) who invited me to collaborate with his team of international scientists, boosting my motivation when I needed it the most. Without him, my work would have taken quite a different shape.

Many thanks go to my friends and fellow PhD students, Szabolcs Sáringer and Gábor Bende. Having company in similar shoes helps to ease the hardships.

Lastly, I would like to thank my wife and baby daughter for cheering me up after each long day of drafting this thesis. I owe them my good spirit.

References

Abari, K., Rácz, Z. Z., & Olaszy, G. (2011). Formant maps in Hungarian vowels—Online data inventory for research, and education. *Twelfth Annual Conference of the International Speech Communication Association*, 1609–1612. https://www.isca-speech.org/archive/pdfs/interspeech_2011/abari11_interspeech.pdf

Ahissar, M., & Hochstein, S. (2004). The reverse hierarchy theory of visual perceptual learning. *Trends in Cognitive Sciences*, *8*(10), 457–464. https://doi.org/10.1016/j.tics.2004.08.011

Angele, B., Baciero, A., Gómez, P., & Perea, M. (2022). Does online masked priming pass the test? The effects of prime exposure duration on masked identity priming. *Behavior Research Methods*. https://doi.org/10.3758/s13428-021-01742-y

Arciuli, J., & Simpson, I. C. (2012). Statistical Learning Is Related to Reading Ability in Children and Adults. *Cognitive Science*, *36*(2), 286–304. https://doi.org/10.1111/j.1551-6709.2011.01200.x

Arrington, C. N., Ossowski, A. E., Baig, H., Persichetti, E., & Morris, R. (2023). The Impact of Transcranial Magnetic Stimulation on Reading Processes: A Systematic Review. *Neuropsychology Review*, *33*(1), 255–277. https://doi.org/10.1007/s11065-022-09534-7

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*, *67*(1), Article 1. https://doi.org/10.18637/jss.v067.i01

Benyhe, A., & Csibri, P. (2021). Can rotated words be processed automatically? Evidence from rotated repetition priming. *Memory & Cognition*, *49*(6), 1163–1171. https://doi.org/10.3758/s13421-021-01147-4

Benyhe, A., Labusch, M., & Perea, M. (2023). Just a mark: Diacritic function does not play a role in the early stages of visual word recognition. *Psychonomic Bulletin & Review*. https://doi.org/10.3758/s13423-022-02244-4

Borgwaldt, S. R., Hellwig, F. M., & de Groot, A. M. B. (2005). Word-initial entropy in five languages: Letter to sound, and sound to letter. *Written Language & Literacy*, *7*(2), 165–184. https://doi.org/10.1075/wll.7.2.03bor

Bouhali, F., Thiebaut de Schotten, M., Pinel, P., Poupon, C., Mangin, J.-F., Dehaene, S., & Cohen, L. (2014). Anatomical Connections of the Visual Word Form Area. *The Journal of Neuroscience*, *34*(46), 15402. https://doi.org/10.1523/JNEUROSCI.4918-13.2014

Brysbaert, M., & Stevens, M. (2018). Power analysis and effect size in mixed effects models: A tutorial. *Journal of Cognition*, *1*(1), Article 1. https://doi.org/10.5334/joc.10

Büchel, C., Price, C., & Friston, K. (1998). A multimodal language region in the ventral visual pathway. *Nature*, *394*(6690), Article 6690. https://doi.org/10.1038/28389

Bürkner, P.-C. (2017). brms: An R package for Bayesian multilevel models using Stan. *Journal of Statistical Software*, *80*, 1–28. https://doi.org/10.18637/jss.v080.i01

Changizi, M. A., Zhang, Q., Ye, H., & Shimojo, S. (2006). The Structures of Letters and Symbols throughout Human History Are Selected to Match Those Found in Objects in Natural Scenes. *The American Naturalist*, *167*(5), E117–E139. https://doi.org/10.1086/502806

Chen, L., Wassermann, D., Abrams, D. A., Kochalka, J., Gallardo-Diez, G., & Menon, V. (2019). The visual word form area (VWFA) is part of both language and attention circuitry. *Nature Communications*, *10*(1), Article 1. https://doi.org/10.1038/s41467-019-13634-z

Chetail, F., & Boursain, E. (2019). Shared or separated representations for letters with diacritics? *Psychonomic Bulletin & Review*, *26*(1), 347–352. https://doi.org/10.3758/s13423-018-1503-0

Cohen, L., Dehaene, S., Vinckier, F., Jobert, A., & Montavont, A. (2008). Reading normal and degraded words: Contribution of the dorsal and ventral visual pathways. *NeuroImage*, *40*(1), 353–366. https://doi.org/10.1016/j.neuroimage.2007.11.036

Cohen, L., Lehéricy, S., Chochon, F., Lemer, C., Rivaud, S., & Dehaene, S. (2002). Language-specific tuning of visual cortex? Functional properties of the Visual Word Form Area. *Brain*, *125*(5), 1054–1069. https://doi.org/10.1093/brain/awf094

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. https://doi.org/10.1037/0033-295X.108.1.204

Corballis, M. C., & Armstrong, C. (2007). Repetition blindness is orientation blind. *Memory & Cognition*, *35*(2), 372–380. https://doi.org/10.3758/BF03193458

Dehaene, S., & Cohen, L. (2007). Cultural Recycling of Cortical Maps. *Neuron*, *56*(2), 384–398. https://doi.org/10.1016/j.neuron.2007.10.004

Dehaene, S., Cohen, L., Sigman, M., & Vinckier, F. (2005). The neural code for written words: A proposal. *Trends in Cognitive Sciences*, *9*(7), 335–341. https://doi.org/10.1016/j.tics.2005.05.004

Dehaene, S., Naccache, L., Cohen, L., Bihan, D. L., Mangin, J.-F., Poline, J.-B., & Rivière, D. (2001). Cerebral mechanisms of word masking and unconscious repetition priming. *Nature Neuroscience*, *4*(7), Article 7. https://doi.org/10.1038/89551

Dejerine, J. (1892). Contribution à l'étude anatomopathologique et clinique des différents variétés de cécité verbale. *Mémoires de La Société de Biologie*, *4*, 61–90.

DiCarlo, J. J., Zoccolan, D., & Rust, N. C. (2012). How Does the Brain Solve Visual Object Recognition? *Neuron*, *73*(3), 415–434. https://doi.org/10.1016/j.neuron.2012.01.010

Dyde, R. T., Jenkin, M. R., & Harris, L. R. (2006). The subjective visual vertical and the perceptual upright. *Experimental Brain Research*, *173*(4), 612–622. https://doi.org/10.1007/s00221-006-0405-y

Fernández-López, M., Mirault, J., Grainger, J., & Perea, M. (2021). How resilient is reading to letter rotations? A parafoveal preview investigation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *47*(12), 2029–2042. https://doi.org/10.1037/xlm0000999

Fernández-López, M., Perea, M., & Vergara-Martínez, M. (2022). On the time course of the tolerance of letter detectors to rotations: A masked priming ERP investigation. *Neuropsychologia*, *172*, 108259. https://doi.org/10.1016/j.neuropsychologia.2022.108259

Forster, K. I., Davis, C., Schoknecht, C., & Carter, R. (1987). Masked priming with graphemically related forms: Repetition or partial activation? *The Quarterly Journal of Experimental Psychology Section A*, *39*(2), 211–251. https://doi.org/10.1080/14640748708401785

Gagl, B., Gregorova, K., Golch, J., Hawelka, S., Sassenhagen, J., Tavano, A., Poeppel, D., & Fiebach, C. J. (2022). Eye movements during text reading align with the rate of speech production. *Nature Human Behaviour*, *6*(3), Article 3. https://doi.org/10.1038/s41562-021-01215-4

Gomez, P., Ratcliff, R., & Perea, M. (2008). The overlap model: A model of letter position coding. *Psychological Review*, *115*, 577–600. https://doi.org/10.1037/a0012667

Grainger, J. (2018). Orthographic processing: A 'mid-level' vision of reading: The 44th Sir Frederic Bartlett Lecture. *Quarterly Journal of Experimental Psychology*, *71*(2), 335–359. https://doi.org/10.1080/17470218.2017.1314515

Grainger, J., Dufau, S., Montant, M., Ziegler, J. C., & Fagot, J. (2012). Orthographic Processing in Baboons (Papio papio). *Science*, *336*(6078), 245–248. https://doi.org/10.1126/science.1218152

Grainger, J., & Jacobs, A. M. (1999). Temporal Integration of Information in Orthographic Priming. *Visual Cognition*, *6*(3–4), 461–492. https://doi.org/10.1080/135062899395064

Grainger, J., & Whitney, C. (2004). Does the huamn mnid raed wrods as a wlohe? *Trends in Cognitive Sciences*, *8*, 58–59. https://doi.org/10.1016/j.tics.2003.11.006

Hannagan, T., Agrawal, A., Cohen, L., & Dehaene, S. (2021). Emergence of a compositional neural code for written words: Recycling of a convolutional neural network for reading. *Proceedings of the National Academy of Sciences of the United States of America*, *118*(46), e2104779118. https://doi.org/10.1073/pnas.2104779118

Harris, I. M., & Dux, P. E. (2005). Orientation-invariant object recognition: Evidence from repetition blindness. *Cognition*, *95*(1), 73–93. https://doi.org/10.1016/j.cognition.2004.02.006

Harris, I. M., Dux, P. E., Benito, C. T., & Leek, E. C. (2008). Orientation Sensitivity at Different Stages of Object Processing: Evidence from Repetition Priming and Naming. *PLOS ONE*, *3*(5), e2256. https://doi.org/10.1371/journal.pone.0002256

Harris, I. M., Harris, J. A., & Caine, D. (2001). Object Orientation Agnosia: A Failure to Find the Axis? *Journal of Cognitive Neuroscience*, *13*(6), 800–812. https://doi.org/10.1162/08989290152541467

Holcomb, P. J., & Grainger, J. (2007). Exploring the temporal dynamics of visual word recognition in the masked repetition priming paradigm using event-related potentials. *Brain Research*, *1180*, 39–58. https://doi.org/10.1016/j.brainres.2007.06.110

Hubel, D. H., & Wiesel, T. N. (1962). Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. *The Journal of Physiology*, *160*(1), 106-154.2.

Hung, C. P., Kreiman, G., Poggio, T., & DiCarlo, J. J. (2005). Fast Readout of Object Identity from Macaque Inferior Temporal Cortex. *Science*, *310*(5749), 863–866. https://doi.org/10.1126/science.1117593

Kinoshita, S., & Kaplan, L. (2008). Priming of Abstract Letter Identities in the Letter Match Task. *Quarterly Journal of Experimental Psychology*, *61*(12), 1873–1885. https://doi.org/10.1080/17470210701781114

Kinoshita, S., Yu, L., Verdonschot, R. G., & Norris, D. (2021). Letter identity and visual similarity in the processing of diacritic letters. *Memory & Cognition*, *49*(4), 815–825. https://doi.org/10.3758/s13421-020-01125-2

Kleiner, M., Brainard, D., & Pelli, D. (2007). *What's new in Psychtoolbox-3?*

Koch, K., McLean, J., Segev, R., Freed, M. A., Berry, M. J., Balasubramanian, V., & Sterling, P. (2006). How Much the Eye Tells the Brain. *Current Biology*, *16*(14), 1428–1434. https://doi.org/10.1016/j.cub.2006.05.056

Koriat, A., & Norman, J. (1985). Reading rotated words. *Journal of Experimental Psychology: Human Perception and Performance*, *11*(4), 490–508. https://doi.org/10.1037/0096-1523.11.4.490

Koriat, A., & Norman, J. (1989). Why is word recognition impaired by disorientation while the identification of single letters is not? *Journal of Experimental Psychology: Human Perception and Performance*, *15*(1), 153.

Kravitz, D. J., Saleem, K. S., Baker, C. I., Ungerleider, L. G., & Mishkin, M. (2013). The ventral visual pathway: An expanded neural framework for the processing of object quality. *Trends in Cognitive Sciences*, *17*(1), 26–49. https://doi.org/10.1016/j.tics.2012.10.011

Lenth, R., Singmann, H., Love, J., Buerkner, P., & Herve, M. (2018). Emmeans: Estimated marginal means, aka least-squares means. *R Package Version*, *1*(1), 3.

Li, X., Huang, L., Yao, P., & Hyönä, J. (2022). Universal and specific reading mechanisms across different writing systems. *Nature Reviews Psychology*, *1*(3), Article 3. https://doi.org/10.1038/s44159-022-00022-6

Lindsay, D. S., Jack, P. C., & Christian, M. A. (1991). Other-race face perception. *Journal of Applied Psychology*, *76*(4), 587–589. https://doi.org/10.1037/0021-9010.76.4.587

Marcet, A., & Perea, M. (2017). Is nevtral NEUTRAL? Visual similarity effects in the early phases of written-word recognition. *Psychonomic Bulletin & Review*, *24*(4), 1180–1185. https://doi.org/10.3758/s13423-016-1180-9

Marcet, A., & Perea, M. (2018). Can I order a burger at rnacdonalds.com? Visual similarity effects of multi-letter combinations at the early stages of word recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *44*(5), 699–706. https://doi.org/10.1037/xlm0000477

McCandliss, B. D., Cohen, L., & Dehaene, S. (2003). The visual word form area: Expertise for reading in the fusiform gyrus. *Trends in Cognitive Sciences*, *7*(7), 293–299. https://doi.org/10.1016/S1364-6613(03)00134-7

Morris, A. L., & Still, M. L. (2012). Orthographic similarity: The case of "reversed anagrams." *Memory & Cognition*, *40*(5), 779–790. https://doi.org/10.3758/s13421-012-0183-7

Mullins, D. A., Whitehouse, H., & Atkinson, Q. D. (2013). The role of writing and recordkeeping in the cultural evolution of human cooperation. *Journal of Economic Behavior & Organization*, *90*, S141–S151. https://doi.org/10.1016/j.jebo.2012.12.017

Nilsson, D.-E. (2021). The Diversity of Eyes and Vision. *Annual Review of Vision Science*, *7*(1), 19–41. https://doi.org/10.1146/annurev-vision-121820-074736

Norris, D. (2006). The Bayesian reader: Explaining word recognition as an optimal Bayesian decision process. *Psychological Review*, *113*(2), 327–357. https://doi.org/10.1037/0033-295X.113.2.327

Norris, D., & Kinoshita, S. (2012). Reading through a noisy channel: Why there's nothing special about the perception of orthography. *Psychological Review*, *119*(3), 517–545. https://doi.org/10.1037/a0028450

Oravecz, C., Váradi, T., & Sass, B. (2014). The Hungarian gigaword corpus. *Proceedings of the Ninth International Conference on Language Resources and Evaluation (LREC'14)*, 1719–1723. http://www.lrec-conf.org/proceedings/lrec2014/pdf/681_Paper.pdf

Pacton, S., Perruchet, P., Fayol, M., & Cleeremans, A. (2001). Implicit learning out of the lab: The case of orthographic regularities. *Journal of Experimental Psychology: General*, *130*(3), 401–426. https://doi.org/10.1037/0096-3445.130.3.401

Peirce, J., Hirst, R., & MacAskill, M. (2022). *Building experiments in PsychoPy*. Sage.

Perea, M., Fernández-López, M., & Marcet, A. (2020). What is the letter é? *Scientific Studies of Reading*, *24*(5), 434–443. https://doi.org/10.1080/10888438.2019.1689570

Perea, M., Gomez, P., & Baciero, A. (2022). Do diacritics entail an early processing cost in the absence of abstract representations? Evidence from masked priming in English. *Language and Speech*. https://doi.org/10.1177/00238309221078321

Perea, M., Hyönä, J., & Marcet, A. (2022). Does vowel harmony affect visual word recognition? Evidence from Finnish. *Journal of Experimental Psychology: Learning, Memory, and Cognition*. https://doi.org/10.1037/xlm0000907

Perea, M., Labusch, M., & Marcet, A. (2022). How are words with diacritical vowels represented in the mental lexicon? Evidence from Spanish and German. *Language, Cognition and Neuroscience*, *0*(0), 1–12. https://doi.org/10.1080/23273798.2021.1985536

Perea, M., Marcet, A., & Fernández-López, M. (2018). Does letter rotation slow down orthographic processing in word recognition? *Psychonomic Bulletin & Review*, *25*(6), 2295–2300.

Perea, M., Vergara-Martínez, M., Marcet, A., Mallouh, R. A., & Fernández-López, M. (2020). When does rotation disrupt letter encoding? Testing the resilience of letter detectors in the initial moments of processing. *Memory & Cognition*, 1–6.

Pinheiro, S., Mota, N. B., Sigman, M., Fernández-Slezak, D., Guerreiro, A., Tófoli, L. F., Cecchi, G., Copelli, M., & Ribeiro, S. (2020). The history of writing reflects the effects of education on discourse structure: Implications for literacy, orality, psychosis and the axial age. *Trends in Neuroscience and Education*, *21*, 100142. https://doi.org/10.1016/j.tine.2020.100142

Price, C. J., & Devlin, J. T. (2003). The myth of the visual word form area. *NeuroImage*, *19*(3), 473–481. https://doi.org/10.1016/S1053-8119(03)00084-3

R Core Team. (2021). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. https://www.R-project.org/

Rajalingham, R., Kar, K., Sanghavi, S., Dehaene, S., & DiCarlo, J. J. (2020). The inferior temporal cortex is a potential cortical precursor of orthographic processing in untrained monkeys. *Nature Communications*, *11*(1), Article 1. https://doi.org/10.1038/s41467-020-17714-3

Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin*, *124*(3), 372–422. https://doi.org/10.1037/0033-2909.124.3.372

Reich, L., Szwed, M., Cohen, L., & Amedi, A. (2011). A Ventral Visual Stream Reading Center Independent of Visual Experience. *Current Biology*, *21*(5), 363–368. https://doi.org/10.1016/j.cub.2011.01.040

Reichle, E. D. (2021). Models of Word Identification. In *Computational Models of Reading: A Handbook*. Oxford University Press.

Rey, A., Dufau, S., Massol, S., & Grainger, J. (2009). Testing computational models of letter perception with item-level event-related potentials. *Cognitive Neuropsychology*, *26*(1), 7–22. https://doi.org/10.1080/09541440802176300

Risko, E. F., Medimorec, S., Chisholm, J., & Kingstone, A. (2014). Rotating With Rotated Text: A Natural Behavior Approach to Investigating Cognitive Offloading. *Cognitive Science*, *38*(3), 537–564. https://doi.org/10.1111/cogs.12087

Schmolesky, M. T., Wang, Y., Hanes, D. P., Thompson, K. G., Leutgeb, S., Schall, J. D., & Leventhal, A. G. (1998). Signal Timing Across the Macaque Visual System. *Journal of Neurophysiology*, *79*(6), 3272–3278. https://doi.org/10.1152/jn.1998.79.6.3272

Seidenberg, M. S., & McClelland, J. L. (1989). A distributed, developmental model of word recognition and naming. *Psychological Review*, *96*, 523–568. https://doi.org/10.1037/0033-295X.96.4.523

Snell, J., & Theeuwes, J. (2020). A story about statistical learning in a story: Regularities impact eye movements during book reading. *Journal of Memory and Language*, *113*, 104127. https://doi.org/10.1016/j.jml.2020.104127

Snell, J., van Leipsig, S., Grainger, J., & Meeter, M. (2018). OB1-reader: A model of word recognition and eye movements in text reading. *Psychological Review*, *125*(6), 969–984. https://doi.org/10.1037/rev0000119

Stevens, W. D., Kravitz, D. J., Peng, C. S., Tessler, M. H., & Martin, A. (2017). Privileged Functional Connectivity between the Visual Word Form Area and the Language System. *Journal of Neuroscience*, *37*(21), 5288–5297. https://doi.org/10.1523/JNEUROSCI.0138-17.2017

Vogel, A., Petersen, S., & Schlaggar, B. (2014). The VWFA: It’s not just for words anymore. *Frontiers in Human Neuroscience*, *8*. https://www.frontiersin.org/articles/10.3389/fnhum.2014.00088

Vogels, R. (2022). More Than the Face: Representations of Bodies in the Inferior Temporal Cortex. *Annual Review of Vision Science*, *8*(1), 383–405. https://doi.org/10.1146/annurev-vision-100720-113429

Whitney, C. (2001). How the brain encodes the order of letters in a printed word: The SERIOL model and selective literature review. *Psychonomic Bulletin & Review*, *8*(2), 221–243. https://doi.org/10.3758/BF03196158

Whitney, C. (2002). An explanation of the length effect for rotated words. *Cognitive Systems Research*, *3*(1), 113–119. https://doi.org/10.1016/S1389-0417(01)00050-X

Wickham, H. (2016). *ggplot2: Elegant graphics for data analysis*. springer.