University of Szeged
Albert Szent-Györgyi Medical School
Doctoral School of Theoretical Medicine

Elusive Aspects of Visual Word Form Processing: Rotations and Diacritics

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

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

1 Introduction

In recent centuries, reading and writing has become a critically important mode of communication among humans. Yet, the time window since the development of writing 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. Crucially, reading-specific functions seem to inhabit the same cortical area in most humans, the visual word form area (VWFA) in the left fusiform gyrus. 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.

Experienced readers (likely exemplified by the reader of this text) are able to decipher written words at a glance and with apparent ease. 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. 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. 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, 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.

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. Starting from the well-known phenomenon of orientation-sensitivity of V1 arising from multiple point like inputs arranged in a row, we can follow the same idea to the emergence of invariant letter representations. In reading, 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. Once the letters of a word are encoded, the word form processing system must also decode their order, as the same set of letters can make up multiple words (e.g., the anagrams 'night' vs. 'thing').

There are several computational models detailing how the brain could solve orthographic processing. Ideally these models should explain phenomena seen in human readers, for instance word identification, word frequency effect or word superiority effect. Most early models focused on these, and neglected the apparent fuzziness of the code, observable when reading words with transposed letters (*you can probalby raed this with esae?*). If the code operated with encoding each letter in a fixed position (e.g., *read = 1:r 2:e 3:a 4:d*), *raed* and *riod* would perform equally as 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. Aiming at this deficiency, multiple models were developed to resolve the problem of letter positions. Such are the Serial Encoding Regulated by Inputs to Oscillations within Letter units (SERIOL) model, the overlap model and the Bayesian reader model. Each explains the letter position effect with a different theoretical background. Importantly, computational models can be tested, as they give quantifiable predictions, which can be compared to experimental data.

We can find some aspects of reading that are rarely captured in computational models and could be useful in their further development. One such element is word orientation. We are surrounded by texts that are rotated from the reference frame of our body (e.g., the spines of books). Although, these may not cause a noticeable difficulty to us, there is evidence for serious cost in reading rotated words. The delay in response times is usually explained by mental rotation, and the LCD model postulates that letter detectors cannot perform with rotations above 45°. There is, however, compelling evidence that letter recognition is highly resilient against rotations, thus the source of reading cost must come from later steps in the processing.

Another elusive feature in reading models is the nature of diacritical letters. Many writing systems use markings that extend the basic Latin alphabet. Some writing systems regard to these as letters in their own right (e.g., the Hungarian *é*) while others see them as modified versions of the base letter (e.g., the Spanish *é*). The question arises, how these letters are coded by the VWFA. There is some evidence that the linguistic function of diacritics modulates the coding status of accented letters. Meanwhile, Bayesian reading models argue that these effects can be accounted for by purely visual features.

2 Aims

No studies so far have presented a detailed view on the psychophysical effects of word orientation on automatic word form 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 previous theory and research, we expected to find a steep change in the presence of rotated priming effects; the LCD model 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.

The encoding status of diacritical letters is still debated. 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, we intended to study the acute accent in Hungarian, a language with a unique combination of ubiquitously used diacritics and very transparent 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’ (first experiment) are closer phonetically, than ‘a’ and ‘á’ (second experiment), we expected that if present, the linguistic pattern would be more pronounced in the latter case.

3 Materials and methods

Rotated word priming (Benyhe & Csibri, 2021)

In this paradigm, the masked priming technique was utilized with backward masking and responses consisting of reading the target word aloud (Figure 1). The primes were either the same or unrelated to the target word and were rotated in various degrees. In the first experiment, we tested a range of different priming durations to modulate the orientational effect, and in the second experiment we used reversed letter order for the prime words to test for the separation of letter and word orientations in word form processing. Fifty-three and fourteen native Hungarians participated in the two experiments, respectively.

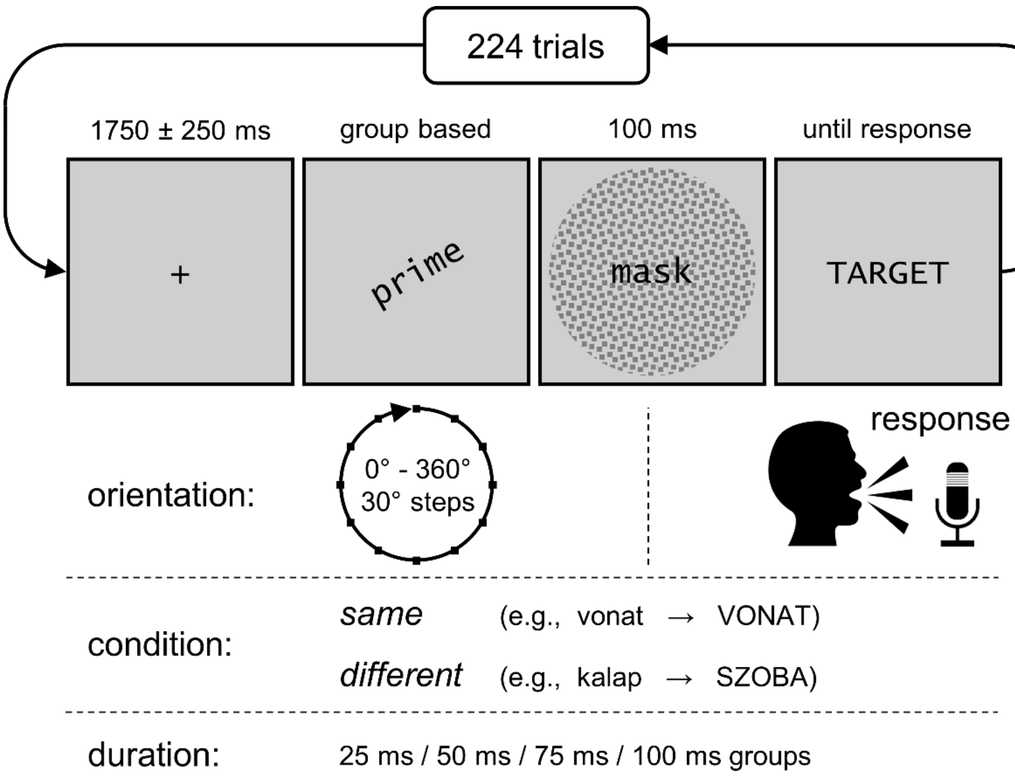
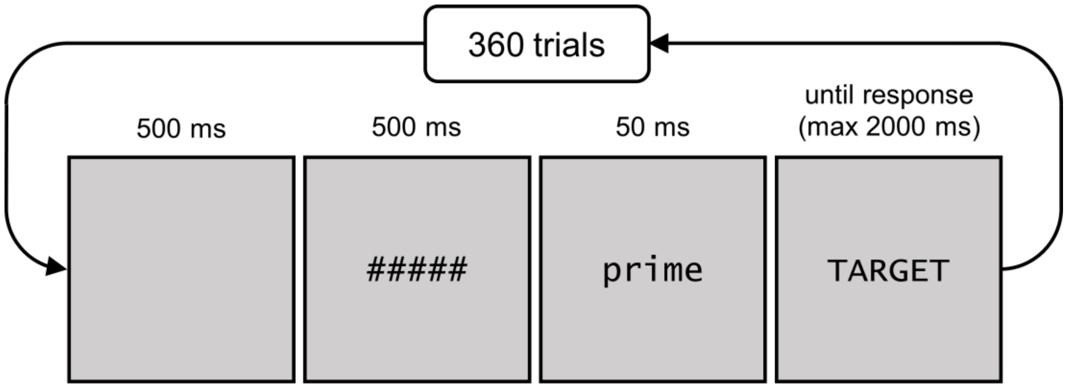


Figure 1 Design of the first experiment of the rotated word paradigm. Each trial starts with a fixation cross of variable interval, followed by the prime stimulus. The prime appears in one of twelve possible orientations and can be of two conditions: a *same* prime is followed by an identical target 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 the second experiment 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).

Diacritic priming (Benyhe et al., 2023)

Two masked priming experiments with lexical decision were designed to evaluate the effects of the acute diacritic used on two sets of vowels to produce two distinct linguistic functions (Figure 2). The long and short versions of Hungarian vowels 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 the first experiment, we modulated diacritic presence on the letters ‘o/ó’ and ‘u/ú’ to change the length of the vowel; in the second experiment, 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’. In both experiments, seventy-two native Hungarian speakers were recruited



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

Figure 2 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.

4 Results

Rotated word priming (Benyhe & Csibri, 2021)

Separate generalized linear mixed effects models were fitted for each priming duration group of the first experiment, 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. Estimated marginal means contrasts detected positive priming effects 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).

With the reversed prime words in the second experiment, there was no significant main effect of prime condition, and there were no apparent positive priming effects in any orientation. The multiple comparisons 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 statistical model

Diacritic priming (Benyhe et al., 2023)

Bayesian mixed effects models were fitted to the RT and accuracy data of both experiments. In the first experiment, the RT model indicated a significant advantage of the *identity* condition (over *similar*), interacting with diacritical, thus the identity advantage was only present for non-diacritical targets (11 ms). Conversely, the *dissimilar* condition had a disadvantage (compared to *similar*), but this was unaffected by diacritics, showing a similar pattern for both diacritical and non-diacritical targets (20 ms and 13 ms, respectively). The accuracy model found no significant effect of the *identity* or *dissimilar* conditions, and their interaction with diacritical status was also insignificant.

In the second experiment (analogous to the previous one), the *identity* condition was once again faster than the *similar* condition, but only for the non-diacritical targets (12 ms). A disadvantage of *dissimilar* primes was present, regardless of diacritical status (11 ms and 16 ms for diacritical and non-diacritical targets, respectively). The *identity* advantage was also present in the accuracy model, with the same interaction pattern (2.2% for non-diacritical targets). The disadvantage of the *dissimilar* condition was, however, absent from the accuracy model.

5 Discussion

Rotated word priming (Benyhe & Csibri, 2021)

In the first experiment, we found that the priming effect was strongly dependent on orientation, which is in contrast with previous findings 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. The range of the rotated priming effect is roughly in line with the LCD model's predictions but is a bit wider and more gradual in its decay. Also, the effect has a moderate time-dependence, as only the 25 ms prime duration had a narrow orientation range, but the longer durations had very similar priming patterns.

When trying to account for rotation effects, most models assume that the delay is caused by mental rotation. In our results, however, the shortness of the prime-target interval, and the masking renders conscious efforts at stimulus normalization improbable. Also, the mixed nature of our design (as opposed to block designs in earlier experiments) 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.

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 would be compatible with modern concepts of the VWFA's role as a modulator of spatial attention.

As letter recognition has been shown to withstand such extreme rotations, we would expect the priming pattern to flip in the second experiment (e.g., in the 180° *same* condition, the letter identities are shared between the prime and the target in each position). As we found no evidence of priming here, we conclude that the resilience of orthographic processing towards rotations is greatest when letter and word orientations agree.

We propose that the difference between unprimed paradigms and our results in the cost of word orientation can be explained by Bayesian principles. 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.

Diacritic priming (Benyhe et al., 2023)

The two experiments yielded strikingly 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*.

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. The fact that the changes in diacritical status produced the same priming pattern in Hungarian as in Spanish and even English studies points out that these effects arise along the visual route. This is in line with the predictions of Bayesian models, 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, 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 earlier, however, that priming effects are case-independent (in other words symmetrical), both for letters with similar and dissimilar features in different cases. 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).

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 or in Spanish. In English, however the representation is expected to be much weaker, as the use of diacritics is less frequent, thus we find weak interference.

General discussion

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. Therefore, the capacities of the reading network will necessarily reflect the challenges it has to solve. 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. Hence, the similarity between diacritical and non-diacritical letters is non-commutable, i.e., asymmetrical. This realization could inspire a new style of modeling, one that builds on the visual variability of prior experience with the written world.

6 Conclusions

- 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°.
- Rotated word priming had a moderate time-dependence, arising for prime durations above 25 ms.
- Our 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.

- We found that similarity priming with omitted or added diacritics was not dependent on the linguistic function of the diacritic, despite clear differences in pronunciation.
- The results in Hungarian revealed the same asymmetry as in other languages: priming occurs with non-diacritical prime and diacritical target, but not in the opposite direction.
- The results 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.