The Effect of Non-word Head Frequency on Letter Position Binding in Word Recognition

Mariam Kabil

S4006720

Department of Psychology, University of Groningen

PSB3E-BT15: Bachelor Thesis

Group number PSB3-BT_2122-|2|a|-|0|4|

Supervisor: dr. Pieter H. de Vries

Second evaluator: dr. W. Robbert van der Mijn

In collaboration with: Melina-Maria Malea, Lukas Schoell, Ruoyu Yu, and Ylse van der Wal.

July 4, 2022

Abstract

This thesis serves as a follow-up in a series of studies investigating position-specific letter recognition based on a conceptual network. To assess letter recall accuracy, an nth-letter task for words and non-words (random string of letters) was utilized. Previous studies have found an unexpected relative peak in letter recall accuracy at the third position in Dutch non-words, that could not be attributed to a manipulation of attention by a centered warning cue on the target position. As upon a visual inspection of the Dutch non-words that elicited the relative peak in performance, it was found that these non-words had quite common head frequencies (i.e., the first three letters in a five-letter sequence made up the beginning of several common Dutch words), the study at hand is investigating the role of head frequency on German (non-) words. The results of the study showed a significant effect for the high head frequency, as it elicited better letter recall accuracy for non-words in comparison with non-words with a low head frequency. Furthermore, the findings of the present study are in support of the serial binding process of the conceptual network, however, some violations of the expected decay pattern need further explorations to fully confirm the processes proposed by the model.

Keywords: word- and letter recognition, conceptual network, serial binding, nth-letter task

The Effect of (Non-) Word Head Frequency on Letter Position Binding in Word Recognition

In an overstimulating world filled with numerous ambiguous stimuli daily, we often try to make sense of these stimuli, by seeking connections. For instance, when presented with ambiguous images or patterns, we often tend to see faces in those – a phenomenon referred to as pareidolia (Pareidolia, 2022). When we read a word, we understand what it means, because that word is connected to a meaning in our brain. Yet, you cna probably stlil raed tihs sentnece, even though it entails words with rearranged inner letters. Can we understand it because we seek connections when reading? Do we read every letter by itself, or do we read words as a whole?

The process of word recognition is a fascinating topic, that has received a lot of attention in the research domain. There are several models that have been developed to address letter and word recognition including the interactive activation model (IAM; McClelland & Rumelhart, 1981), and the SERIOL Model (Whitney 2001). Most of those models are described at a functional level, i.e., they outline how information is processed. However, there is another level of description that refers to the underlying mechanisms that are required for biologically plausible self-organization: the structural level. An example of such a structural model based on the properties of neurons is the conceptual network, which will be the model the author focalizes in this paper. The study conducted in this paper is a follow-up study on the experimental work of previous Bachelor's students (e.g., Mudogo, 2019; Pink, 2019; Schwartzkopf, 2019; Bhouri, 2018; Donelan, 2018; Freericks, 2018), that were based on the conceptual network and investigated the task of reporting a letter at a given position (letter recall accuracy) for a word or a non-word. From now on we will refer to this task as an nth-letter task.

Major Approaches to Word Recognition

One approach to word recognition is the position-specific slot encoding. This approach suggests that we recognize words based on the positions of the letters included in that word. Words are encoded by using a slot for each letter; thus, one letter is always combined with a certain position (Davis & Bowers, 2006). For example, the word shy is coded as S_1, H_2, Y_3. This approach is utilized in the interactive activation model (IAM), that i.a. established the *word superiority effect:* letters are better recognized when they are part of a real word, rather than when part of a non-word (random string of letters), or when they are presented on their own. Position-specific slot encoding can however not account for transposition priming (in which priming for a word is still successful, although the position of the letters is changed). The IAM further falls short of the central task in this thesis, as it is not useful if you want to report a letter at a given position, because there needs to be top-down activation from the word node to the letter nodes **specific** to this word, not just general top-down activation as previously mentioned in other studies (Whittaker, 2019). Additionally, words like "GARDEN" can still be read, when the letters are not in their original position as in "GADREN". In order to address these limitations, another approach can be employed: the open-bigram coding, as for example used in the SERIOL Model (Whitney, 2001). Here, it is proposed that words are coded by means of the letter pairs that make up that word (Grainger & Whitney, 2004). For instance, the word "ACT" can be coded by a set of the following letter pairs [AC, AT, CT]. This type of coding however also does not imply position-specific top-down activation or any kind of top-down activation, and thus also falls short of the central task in this thesis.

The Conceptual Network

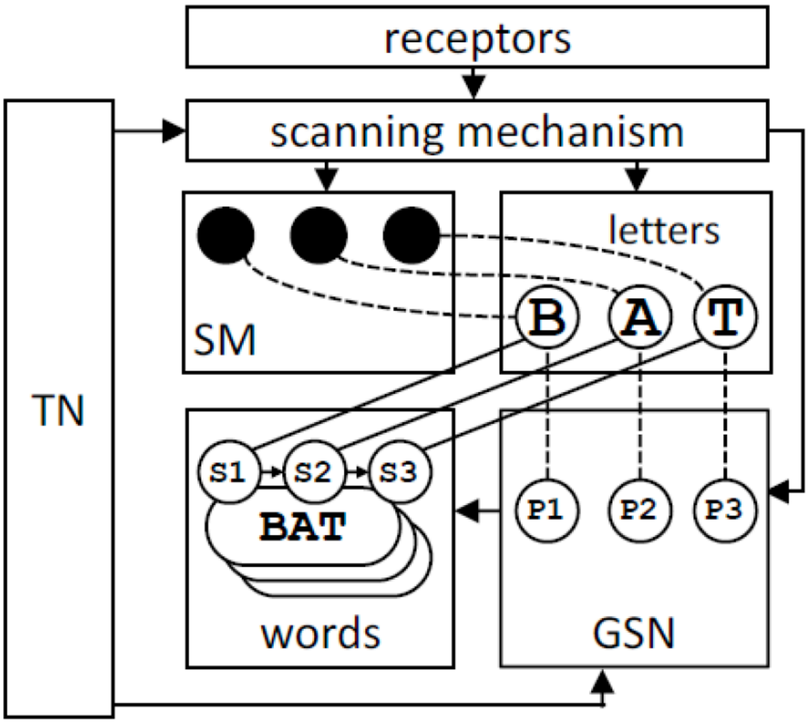
Reading a word requires us to process two distinct features, the identity of the letters included in that word, as well as the position of the letters. If these features were represented jointly in our brain, a plethora of memory traces would be needed to recognize and comprehend words. Thus, the more realistic assumption is that letter identity and letter position have separate representations in the brain. In the process of reading a word, these two features must then become bound, yet how does this connection take place? – A riddle referred to as the binding problem (von der Malsburg, 1999). Both approaches presented above yet still fail to explain how letter identities are exactly bound to specific letter positions at a structural level. A model, that however does offer an explanation to this problem is the conceptual network (de Vries, 2016).

The conceptual network is based on the notion of the Hebb rule that claims that neurons that fire together, wire together (Hebb, 1949); neurons that are activated at the same time form a stronger synaptic connection than neurons that are not. As a consequence of that rule *cell assemblies* are formed. These assemblies are clusters of neurons that have more and stronger connections with each other than with neurons in other clusters, with one assembly corresponding to a memory trace or a concept. As neurons fire when a critical threshold is exceeded, cell assemblies act in the same way. A critical threshold denotes the minimum level of excitation needed, for a cell assembly's activation to grow without further external stimulation. When a cell assembly (CA; and thus, corresponding concept), is activated beyond the critical threshold (CT), we are consciously aware of the concept that was activated, as considered from a functional perspective, it presents itself in working memory. Hence, recognizing a word translates to an activation of the corresponding cell assembly at a level beyond that critical threshold. Yet, when the activation level is below the critical threshold, the memory trace is solely in a priming state, and a word is not recognized consciously.

As outlined above, permanent connections between all possible letter identities and letter positions would lead to a surplus of memory traces, considering that also location plays a role and not only letter are involved but also their order in a sequence, what we call a superposition catastrophe (von der Malsburg, 1999). This calls for a solution of the binding problem by means of a temporary connection (i.e., binding), that is based on two conditions (de Vries, 2016). Firstly, two cell assemblies must be activated at the same time in a conceptual network. Secondly, a subnetwork of that conceptual work must exist, that activates memory traces and represents actual context in which they play a role. When memory traces are activated, spike trains produced by neurons of both cell assemblies are synchronized and this is referred to as *spike resonance*. When both conditions are fulfilled, a single temporary connection emerges based on an already existing pathway. However, in word processing, more than one temporary connection is necessary, as for each letter there are two excitation patterns (for a letter's identity and location), and within a word there are several letters. Thus, the binding of identity and location is assumed to take place serially, and not in parallel, because parallel binding would lead to interference in the binding process. Yet, with serial binding, interference could be circumvented, by one bond occurring rapidly after another (de Vries, 2016). Nodes within the conceptual network representing cell assemblies (or letter identities) can form either permanent connections (represented by solid lines in Figure 1), or temporary connections (represented by dashed lines). For instance, word nodes have a permanent connection to letter nodes (i.e., identity of letters). The location of letters is represented in an inborn neural structure referred to as the spatial map (SM), in which temporary excitation patterns can occur due to external activation of the network. These excitation patterns (of various forms of a letter) are then bound to a memory trace (of a letter identity) in the spatial map, leading to an activation of that letter beyond the critical threshold, making a reader aware of that letter at a certain position. Position nodes are represented as memory traces in a neural structure called the global sequence network (GSN). A subnetwork of the GSN is the local sequence network, that is a structure specific for each word representation. How we then recognize words is a two-fold process. Firstly, a top-down activation from the word node level to letter node level takes place. Secondly, a bottom-up activation from the sensory input on the letter node level occurs.

Figure 1

The Conceptual Network



Note. A visual representation of the conceptual network for letter- and word recognition. Modules are included in the form of rectangles such as the input (receptors, scanning mechanism, a spatial map (SM), a task network (TN), a global sequence network (GSN), and subnetworks for words and letters. Adapted from de Vries, P. H. (2016). Neural binding in letter- and word-recognition. In K. E. Twomey, A. Smith, G. Westermann & P. Monaghan (Eds.), *Neurocomputational models of cognitive development and processing: Proceedings of the 14th neural computation and psychology workshop* (pp. 17- 33). New Jersey: World Scientific, 2016.

Previous Studies

Earlier studies that preceded this experiment and used the conceptual network as their basis can be organized into two groups (Previous Studies 1 and 2). The first group of studies was mainly set out to provide support for the basic biological underlying mechanisms of the conceptual network, namely that the serial binding would lead to neural decay. That is, for a string of 5 letters, the strength of binding will decrease with consecutive letter positions (from letters one to four). Yet, at the last position a relatively strong binding is evident, as there is no position to transfer the activity to, an effect referred to as the *reverberation effect*. In Previous Studies 1, participants had to perform an nth-letter task, in which they were presented with five-letter words and non-words, and consequently had to identify a letter at a given position. They hypothesized that for non-words, a hook-shaped distribution of letter recall would take place. In the experiment, the word nodes influence the letter report of non-words in a top-down, and from left to right process. For instance, a high frequency beginning in non-words would lead to an activation from word nodes to letter nodes. More specifically, if the word BAUEN is to be recognized, first the most left position is primed and all words with the same beginning letter (B) are activated, whereas all other words that do not begin with the letter (B) will decay. Next in the serial binding process, all words with the beginning two letters (BA) are activated and the rest of the words that begin with other letters will decay. The same process will repeat itself for the remaining letters in the non-word. Thus, the letter recall accuracy was predicted to be the highest for the first letter with a decrease in performance in the consecutive letters until the second to last letter, resulting in a hook-shaped distribution. This hook-shaped distribution was also indicated by three studies using English and German (non-)words (Mudogo, 2019; Pink, 2019; Schwartzkopf, 2019), however two studies (Buijsman, 2019; Whittaker, 2019) using Dutch (non-) words unexpectedly found a relative peak in letter recall accuracy at the third position for Dutch non-words. For words it was predicted that the recall accuracy of the given letter would be high for all five positions, due to top-down processing of the target word. The results indeed showed an even distribution of performance across the five positions.

In an attempt to explain these unexpected findings, a second set of experiments were conducted that manipulated an attention condition. It was speculated that perhaps the third position in Dutch non-words had increased attention, and thus higher activity was present at that position, leading to increased excitation levels at the corresponding location in the spatial map and consequently more binding occurred there, causing better recall accuracy for that position. In order to assess this hypothesis, (non-) words were cued by a warning signal that was either distributed (equal attention provided to all positions) or centered (increased attention to one specific position). For words, the prediction was made that a distributed attention, as well as centered attention conditions would have no effect on the performance of the participants. That is, the strong top-down processing of words would render the effects of attention on performance insignificant, as high performance would be expected in either way. This hypothesis was indeed largely supported by the findings of Previous Studies 2 (Bhouri, 2018; Donelan, 2018; Freericks, 2018; Seibel, 2018). Interestingly, for the non-words no effect was predicted by manipulating the attention to be distributed, yet as they used a centered cue in the other condition, a response advantage was expected for the target position, on which the attention was centered. The results showed, that as predicted for the distributed attention condition, a hook-shaped distribution was evident. Yet, in contrary to their second hypothesis, no peak in performance was found at the target positions in the central attention condition.

The Present Study

Thus, this present study, aims at explaining the peak in performance found in Previous Studies 1 in Dutch non-words. With closer inspection, it was found that the Dutch non-words used in these studies (Buijsman, 2019; Whittaker, 2019), had quite common head frequencies (i.e., the first three letters of those non-words were common starters of Dutch words). In order to closely inspect whether that peak in performance is indeed due to the head-frequency, a series of studies is set up to investigate the effect of high head frequencies and, for comparison, of high tail frequencies. A high tail frequency here refers to the last three letters of a non-word that make up the ending of common words.

This study is set out to determine the effect of the head frequency on the letter recall accuracy of German words and non-words, as well as the influence of letter position on the letter recall of those (non-) words in the context of the conceptual model. In order to assess the first effect, a base condition (of (non-)words with low head and low tail frequencies) is compared to (non)-words with a high head and low tail frequency. The other experiments that make up this series of studies include three investigating the role of words with high head frequencies, and two investigating the role of words with high tail frequencies in comparison to a control condition of words with low head and low tail frequencies in three different languages (English, German and Dutch).

Hypothesis 1

For the non-words, we expect higher performance on the nth-letter task for non-words with a high head frequency, in comparison to a low head frequency. Thus, we expect to see a hook-shaped distribution across the letter positions, with a decrease in performance for consecutive positions from position one to position four, and a slight increase in position five due to the reverberation effect. These effects are also expected in the other experiments of the same design using Dutch and English (non-) words.

Hypothesis 2

However, for German words, we expect an evenly distributed letter recall accuracy across all five positions, for both conditions (words with a low head, low tail frequency, and words with a high head, low tail frequency).

Hypotheses of the Twin Experiments within the Experimental Group

In the twin experiment within this experimental group (from now on referred to as experiments of Design 1) examining the effects of a high tail frequency (in low head, high tail non-words), we similarly expect an equal letter recall accuracy across positions in words. In non-words, however, we predict to see a hook-shaped distribution, that is very similar for both frequency conditions (low head, low tail non-words, and low head, high tail non-words).

Method

Participants

The convenience sample used was gathered by means of the platform Prolific (https://prolific.co/). The total sample size comprised 49 participants (61.2% female, M_{age} = 21.4, SD_{age} = 3.5, age range 18-30). All participants spoke German as their first language and received a monetary reward of £ 2,00 according to the platform standards. The mean duration of the experiment was 14.2 minutes (*SD* = 5.6). Participation in the experiment was voluntary, and all participants signed informed consent prior to it. The study was approved by the Ethics Committee of the Faculty of Behavioural and Social Sciences, Department of Psychology, at the University of Groningen.

Design

The present experiment was set up in the format of a 2x2x5 design with three independent variables: word type (two levels: words and non-words), frequency (two levels: low head, low tail and high head, low tail), and letter position (five levels: for each of the five letters in a five-letter sequence). In total, these variables amounted to 20 conditions. The dependent variable was the accuracy of letter recollection. This is given by a proportion of correctly recalled letters per condition in the performance trials.

Stimuli

The stimuli used in this study were word and non-word pairs. The paired words were retrieved from the CELEX Centre for Lexical Information (2001) database. Each word pair only differed in one letter in one position (BAUEN, HAUEN). The position where the letter in a pair varies was also the target position the participant had to remember. This was done to account for the *word substitution effect* (Reicher, 1969). For instance, participants might be able to guess the target letter by just its surroundings. Thus, the word pair creates ambiguity where the exact letter has to be correctly remembered. Therefore, the letter itself becomes more important and a participant relies more on the letter than its surroundings.

The non-words were constructed by rearranging the letters of the corresponding words (e.g., BAUEN is rearranged to BEUAN). The same letters were used here in order to preserve similar letter frequencies and thus also the letters' physical properties for the non-words. This way, when we observe changes in letter recall accuracy, we are able to attribute these changes to a letter position and not to the frequency of a letter or its physical properties. Now, words and non-words share letters with the same frequency.

The non-words were constructed in a manner to have a high head frequency and a low tail frequency. As explained in the design section, above, a high head frequency refers to the first three letters that form the beginning of relatively common words. Whereas a low tail frequency refers to the last three letters that do not form the ending of relatively common words. Common words in this context are defined as having a word frequency of >7 per million words.

While constructing the non-words, we could not always keep the target letters at the same position it had in the words, in contrast to the previous studies mentioned above. As this study examines the effect of a non-word's head frequency, the non-words had to be constructed in a way to ensure that frequency. Hence why, the target letter in the non-word was not held at the same position of the word, to guarantee said frequency.

The target letter is the letter that the participant should report. If the second letter was asked to be reported, then this letter was the target letter in the corresponding word or non-word (e.g., B in ABRUF). The participants know which letter was the target letter because below the mask (#@@@@@#), a number was shown, 1 till 5, which corresponds to the letter in the word (as shown in Figure 2).

The stimuli (words and non-words) that were selected and generated were too many to present to one participant, because of the time limitation of the online format. Thus, we created lists of comparable frequencies, so that a subset of participants was presented with a subset of the (non-) words. For each of the ten non-word conditions (non-words x 5 letter positions x 2 frequency) of the experiment, four compatible lists of six non-words were constructed. For the word conditions, 2 lists were generated. Per participant, one list out of four or two was randomly selected for each of the 20 conditions. This way, when we find the same results across all lists, the results cannot be attributed to the specific words or non-words.

Procedure

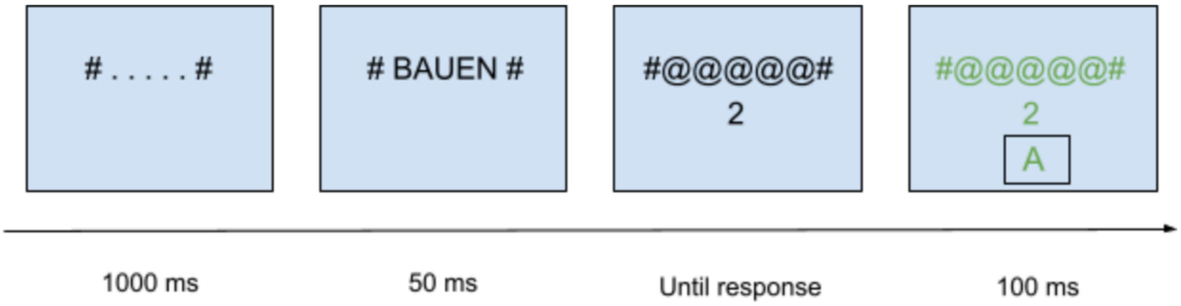
This experiment was conducted in an entirely online setting, in a location of the participants' choosing, and thus its duration had to be limited. Participants were instructed to go to a quiet surrounding with no distractions and complete the experiment on a laptop or PC. Before the experiment started, the participants gave their informed consent after receiving information about their rights and the nature of the research via Qualtrics (Qualtrics, Provo, UT). They were informed that their age, gender, and performance would be collected and that they could contact the principal investigator via mail, should there be any questions.

Next, the participants received the task instructions. Upon the completion of these actions, participants were redirected to a website running a program that was created with the OSWeb tool version 1.4.4 of the program OpenSesame version 3.1 (Mathôt et al., 2012) where the experiment itself took place. The whole experiment included two training blocks and three experimental blocks. Each training block consisted of ten trials. Each experimental block included 40 trials. In total, each participant was presented with 120 (non-) words.

In each trial, four stages followed each other (see Figure 2). The first screen the participants saw was blank. After 1000ms a preparation signal (#.....#) appeared on the screen which lasted for 500ms. The preparation signal indicated where the letters of the stimulus would appear. After the preparation signal, the real target appeared for 50ms (#BAUEN#). Finally, a mask (#@@@@@#) replaced the target. Below the mask, a number would show up, and the corresponding letter of the (non-) word should be filled in by the participant. This screen was there until the participant filled in a letter and pressed enter. In all situations, the dots, the letters, and the '@' were surrounded by '#', in order to ensure consistency on every screen. The stimuli were presented with two hashtags surrounding them, one before the first and one after the last letter because our visual processing is very sensitive to contrasts. That is, the first and last letter's salience (in BAUEN, B, and N) would be too prominent, and would have an unwanted advantage in contrast to the middle letter positions resulting in a higher recall accuracy for positions 1 and 5. Yet, as we are interested in the relationship of serial binding, we add symbols around the first and last letter positions, to accurately test their recall according to the conceptual network.

In the training blocks, after each (non-) word the participant got feedback about the letter they filled in. If it was the right letter the whole screen turned green (see fourth screen, Figure 2), if it was the wrong letter the screen turned red. In this way, participants could get familiarized with the task and get the highest score possible in the experimental block without being confused over the task itself. After each training block the participant received general feedback in a percentage, which reflected the accuracy of the recollection of the letters. In the experimental blocks, only the general feedback was given at the end of a block.

F*igure 2. An example of the sequence of screens in a training block, with a correct answer.*



The stimuli in the practice and performance trials were presented centrally on the computer screen in the font size of 24p. This shows an increase in stimuli size by 6p from previous studies, referred to earlier, as previous studies were carried out on bigger screens in a laboratory with lower resolution. Therefore, we increased the letter size due to the smaller size of the laptops commonly used by the participants in an online setting. At the end of the experiment, the participants received their monetary reward.

Analysis

The experiment used a within-subjects design, as every participant was tested for all experimental conditions. We expect a three-way interaction between the three independent variables. Consequently, the data was analyzed using a three-factor Repeated measures (RM) ANOVA in IBM SPSS Statistics, Version 27.0. Before carrying out the analysis, the raw data was prepared by aggregating and restructuring it in SPSS (Lacroix & Giguère, 2006). In detail, the three independent variables (word type, letter position, and head frequency) were examined for both main- and interaction effects.

Subsequently, four one-way RM ANOVAs with the factor letter position were run, in order to check whether in each of the four sets of five positions the means per letter position follow the expected pattern. This is necessary to get an insight into the expected three-way interaction of the three independent variables. Additionally, pairwise comparisons with Bonferroni corrections were conducted for all RM ANOVAs to assess the distribution of the letter recall accuracy across the experimental conditions.

Results

Based on the theory of the conceptual network, we expected a three-way interaction between the three independent variables. For the non-words, we expected a general difference between the frequency conditions, as such that non-words with a low head frequency will indicate a steeper hook-shaped decline between consecutive letter positions (except for an increase in the last position) in comparison to non-words with a high head frequency. However, for the words, we predicted that there will be no significant difference between the letter positions in both frequency conditions. For the usage of RM ANOVA, the necessary assumptions of normality, sphericity, and independence were checked for. All assumptions held, apart from the assumption of sphericity in some conditions, that will become evident in the tables below, for which Greenhouse-Geisser corrections were made.

The three-factor RM ANOVA analysis show a significant three-way interaction between the three independent variables, $F(3.4,163.41) = 27.9$, $p < .001$, $\eta^2 = .37$, confirming our prediction. All main and interaction effects of all independent variables were significant with $p < .001$ and are shown in Table 1 below. Notably, the interaction effect between word type and position was significant, *F*(2.72, 130.36) = 13.6, $p < .001$, $\eta^2 = .22$, indicating a difference in letter recall accuracy for words vs. non-words across the five positions.

Table 1

Three-Way Repeated Measures ANOVA Analysis

Source	*df*	*F*	*p*	η^2
Frequency	1; 48	78.8	< .001*	.62
Position	4; 192	109.28	< .001*	.7
Word Type	1; 48	298.21	< .001*	.86
Frequency * Position	4; 192	43.33	< .001*	.47
Frequency * Word Type	1; 48	114.47	< .001*	.71
Position * Word Type	2.72^a; 130.36	13.6^a	< .001*^{a}	.22^a
Frequency * Position * Word Type	3.4^a; 163.41	27.9^a	< .001*^{a}	.37^a

Note. This table demonstrates the results of the three-factor RM ANOVA, with the factors word type (2 levels), letter position (5 levels) and frequency (2 levels). Significant values are marked with a (*), using a significance level of ≤ .05.

Note. [a] Refers to values computed using a Greenhouse-Geisser correction.

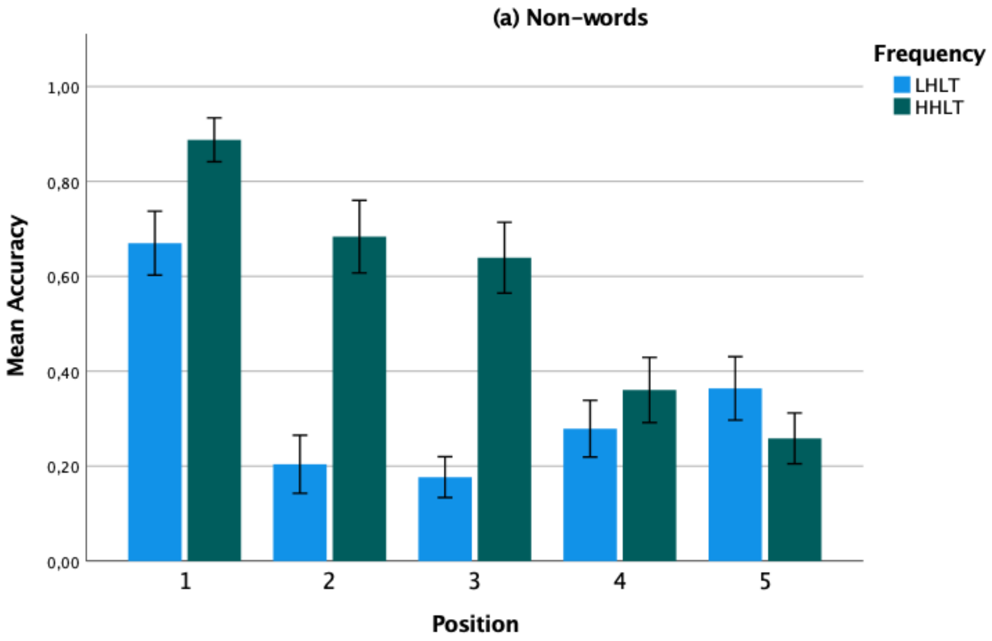
Non-words

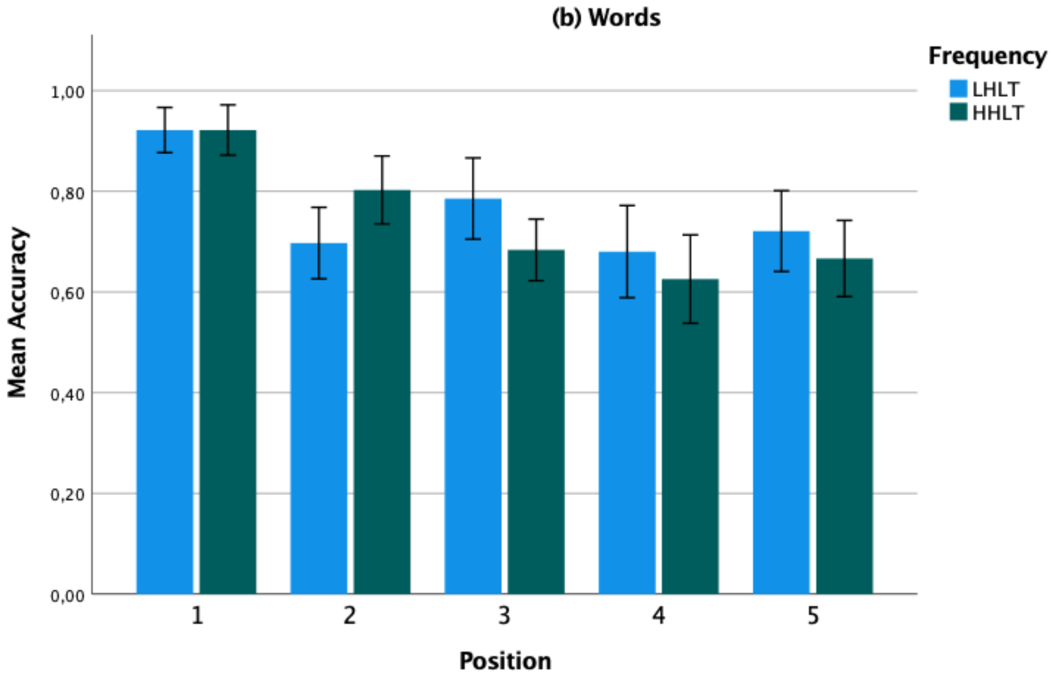
Next, to verify whether the hypotheses on non-words and words actually stand, we need to check whether the pattern of the means over positions follows the expected decay in non-words and shows equal performance across positions in words. Therefore, four one-way RM ANOVAs were computed, alongside pairwise comparisons with Bonferroni corrections.

We predicted a decrease in letter recall accuracy from P1 (position one) to the second to last position (P4), and an increase from P4 to P5 for low head, low tail (LHLT) non-words, resulting in a hook-shaped distribution. As predicted, the results indeed showed a similar distribution for LHLT non-words (see Figure 3a). The pairwise comparisons for LHLT non-words indicated a significant decrease from P1 (*M* = .67, 95% CI [0.61; 0.74]) to P2 (*M* = .21, 95% CI [0.14; 0.27]) and from P1 to P3 (*M* = .18, 95% CI [0.13; 0.22]). However, there was no significant difference between P2 and P3, despite a predicted decrease from P2 to P3. Furthermore, a statistically significant increase from P3 to P4 (*M* = .28, 95% CI [0.22; 0.34]) was found, yet no significant increase from P4 to P5 (*M* = .37, 95% CI [0.3; 0.43]) was apparent, in contrast to what we had expected. Position five was significantly higher than positions two and three. The highest p-value for significance was .017, whereas the lowest p-value for insignificance was .406.

Figure 3

Mean Letter Recall Accuracy for Low Head and High Head Frequency Words and Non-words across all five Letter Positions





Note. The results for non-words (a), and words (b), with the five letter positions on the x-axis and the average letter recall accuracy on the y-axis. The blue bars represent low head and low tail (LHLT) frequency, whereas the green bars show high head and low tail (HHLT) frequency (non-) words. Error bars are indicated with a confidence interval of 95% around the mean.

In regard to high head, low tail (HHLT) non-words, we similarly predicted a hook-shaped distribution with a decrease that is less steep than for LHLT non-words. This shape seems to be partly present (see Graph A, Figure 3), for positions one throughout 4, with the exception of P5 (decrease from P4 to P5, instead of an expected increase). The pairwise comparisons for HHLT non-words indicated a significant decrease from P1 (*M* =.89, 95% CI [0.84; 0.93]) to P2 (*M* =.68, 95% CI [0.61; 0.76]). No significant difference between P2 and P3 (*M* =.64, 95% CI [0.57; 0.71]) was evident. Further, P4 (*M* =.36, 95% CI [0.29; 0.43]) was significantly lower than P3, as well as P5 (*M* =.26, 95% CI [0.21; 0.31]) in comparison to P4.

As predicted, the letter recall accuracy was significantly the highest in P1, second highest in P2, and further decreased from P3 to P4. Unexpectedly, we found no difference between P2 and P3, although a decrease in consecutive positions was predicted. Additionally, and contrary to our predictions, there was no increase from P4 to P5, and P5 significantly displayed the lowest recall accuracy of letters among all positions. The highest p-value for significance was .046, whereas the lowest (and only) p-value for insignificance was 1.0.

Words

For the LHLT words, we expected a higher letter recall accuracy than non-words, yet equal performance over the five positions. However, the pairwise comparisons showed partly significant differences between the letter recall accuracy across the letter positions. For instance, P1 (*M* =.92, 95% CI [0.88; 0.97]) was significantly higher than the other four positions. There were no significant differences between P2 (*M* =.7, 95% CI [0.63; 0.77]) and P3(*M* =.79, 95% CI [0.71; 0.87]), P4 (*M* =.68, 95% CI [0.59; 0.77]) and P5 (*M* =.72, 95% CI [0.64; 0.8]). Additionally, P3 was significantly higher than P4. Thus, letter recall accuracy appears to be best for the first position, in comparison to positions two to five. Further, letter recall accuracy was slightly better for position three in comparison to positions two, four and five. Thus, in contrast to our hypothesis, letter recall accuracy is not equal across positions. The highest p-value for significance was .005, whereas the lowest p-value for insignificance was .133.

Similar predictions were made for the HHLT words, such as that we expected a higher letter recall accuracy than non-words, yet equal performance across the letter positions. However, pairwise comparisons depicted here too partly significant differences between letter recall accuracy across the letter positions. The letter recall accuracy for P1 (*M* =.92, 95% CI [0.87; 0.97]) was significantly higher than the other positions. Similarly, P2 (*M* =.8, 95% CI

[0.74; 0.87]) was significantly higher than P3 (*M* =.68, 95% CI [0.62; 0.75]), P4 (*M* =.63, 95% CI [0.54; 0.71]) and P5 (*M* =.67, 95% CI [0.59; 0.74]). Yet, there were no significant differences between positions three, four and five. These results show, in contrary to our hypothesis, that the letter recall performance is *not* equal across all positions. The highest p-value for significance was .002, whereas the lowest p-value for insignificance was .843.

Based on the patterns of the means presented above, it becomes clear that the factor frequency has a different effect on the letter recall accuracy for the five positions in non-words than for words. This is reflected in a hook-shaped distribution for non-words in contrast to almost a flat distribution for words. Thus, an interaction effect between word-type and frequency is present. Additionally, an interaction effect was evident between word-type and letter position, as the letter recall accuracy across positions was different in non-words in comparison to words. Further, there was an interaction effect between frequency and letter position, as the pattern of letter recall accuracy differed from one position to another, in both conditions: HHLT and LHLT (non-) words. Thus, in line with our hypothesis, a significant three-way interaction exists between the three independent variables.

Discussion

The present study was set out to determine the effect of the head frequency on the letter recall accuracy for German words and non-words, as well as the influence of letter position on the letter recall of those (non-) words in the context of the conceptual model. Two main hypotheses were proposed considering the letter recall accuracy for non-words in comparison to words. In the first section of this discussion, the hypothesis for non-words in the present study will be discussed, followed by an analysis of the results for non-words in the other studies of this series. The second section will follow the same outline yet relating to the hypotheses for words.

The first main hypotheses regarded the non-words, with hypothesis 2a relating to LHLT non-words, and 2b relating to HHLT non-words. Hypothesis 2a proposed a decrease in letter recall accuracy in consecutive positions from P1 to P4, as with each position the temporary bindings between position and identity become weaker due to the activation decay, and an increase in P5 due to the reverberation effect, referred to earlier in this thesis. According to the conceptual network, non-words are not easily recognized as words, as there are no pre-existing mental representations for these non-words, and thus the excitation level of a word node cannot exceed the critical threshold rendering a non-word recognizable. The results of this study demonstrate partial support for the expected decay pattern, as the letter recall accuracy decreased following the first position. However, the expected decrease from P2 to P3 was not found, and there was no significant increase from P4 to P5. Nevertheless, P5 denoted a relative, significant increase in the letter recall accuracy in comparison to positions two and three. However, Malea (in preparation) and van der Wal (in preparation) detected an increase in the last position in Dutch non-words, supporting our predictions.

Hypothesis 2b proposed a similar hook-shaped distribution with a decrease in consecutive positions until the second to last position and an increase in the last position. However, the performance on the nth-letter task was expected to be significantly higher for HHLT non-words in comparison to LHLT non-words. The results of the present study indeed showed a higher performance on the task at hand for HHLT non-words in comparison to the LHLT non-words, providing evidence for our hypothesis. Additionally, in line with the predictions, there was a significant decrease from P1 to P2, and from P3 to P4. However, there were minor violations of the expected decay as the results did not indicate a decrease from P2 to P3, neither an increase in last position. The violation of the expected decay pattern in the last position (in both frequency conditions) could be explained by a participant's expectations. For instance, if the beginning of the non-word formed the beginning of common words, a participant might have expected a different end, due to a top-down activation for another word. As all participants in this study were German native speakers, this is a likely explanation for the error in position five, due to a strong top-down activation of real words. Malea (in preparation) similarly failed to find an increase in the last position in English HHLT non-words, which might be due to the reasons elaborated on earlier, or that might have been influenced by the small number of native speakers of the experiment's participants. Van der Wal (in preparation) however found the expected increase in the last position, further supporting the conceptual network's decay pattern.

This study was embedded in an experimental group, that i.a. examined the effect of a high tail frequency via the usage of two frequency conditions: low head, low tail, and low-head, high tail (non-) words. For both non-words frequency conditions, the same hook-shaped distribution of accuracy across letter positions was expected, as described above. Thus, no difference between conditions was predicted, as the tails of non-words are unlikely to elicit top-down activation for existing words. Interestingly, Schoell (in preparation) indeed found that distribution, however, a high tail seemed to elicit higher performance on the last four positions in comparison to a low tail. These findings raise the question, of whether a high tail can elicit top-down activation for words. Further research focused on that question must be conducted, to clarify the connection between these results and the conceptual network

For the second main hypothesis, we expected an evenly distributed letter recall accuracy across all five positions in German words, for both frequency conditions (LHLT and HHLT words). According to the conceptual network, the retrieval of target letters is facilitated due to the position-specific top-down activation that occurs from word nodes to letter nodes and due to the pre-existing mental representations of words. The results showed that performance on the nth-letter task was not equal across positions, disconfirming our hypothesis. For instance, in both frequency conditions, the first position was significantly higher than the remaining four positions. Additionally, the accuracy for the third position in LHLT words was significantly higher than the fourth position and the second position in HHLT words was significantly higher than P3, P4, and P5. These unexpected results, and especially the peak in performance for P1, could be explained by the noise created by the different (non-) word lists used. Each list consisted only of six words, which amounts to a relatively small sample size of words. Furthermore, despite all participants being native German speakers, they might have not recognized certain words, if the participant's concentration was otherwise occupied or fatigued or if they simply did not know the word. Thus, an experienced reader might have an unwanted advantage of recognizing words quicker than an inexperienced reader. Additionally, as the words were presented for a short period of time (50 ms), the word nodes might not have become activated beyond the critical threshold.

The absence of an effect of frequency on the letter recall regarding their position in a word was also expected in the other studies of this series, and thus we predicted an equal performance across letter positions across all languages and designs. Van der Wal (in preparation) found evidence in support of this hypothesis, as her results for LHLT words showed no difference in letter recall accuracy across positions. However, the hypothesis was only partly supported as e.g., van der Wal's (in preparation) findings on HHLT Dutch words demonstrated a plunge in performance for the fourth letter position. Similarly, Malea (in preparation) found a significant decrease in position three in comparison to the other positions in English HHLT words, as well as Schoell (in preparation) who found a drop in positions three and four for German LHHT words. Taken together, the general letter recall accuracy for words was higher than for non-words and followed roughly the same trend except for the minor violations mentioned above.

In conclusion, the findings presented above show a significant effect of the factor frequency on the letter recall of participants. Thus, reading does not appear to be influenced by a manipulation of attention (Bhouri, 2018; Donelan, 2018; Freericks, 2018; Seibel, 2018), however frequency is a likely explanation for the W-pattern (peak in P3) found by previous studies (Whittaker, 2019). In support of this claim is the finding that the third position of HHLT non-words was significantly higher than the second position in LHLT non-words in all three languages of the same design.

Despite the unexpected differences in letter recall accuracy across letter positions in words, and some violations of the expected decay pattern in the non-words, there is support for the serial binding of letter identity and position, as proposed by the conceptual network. This can be concluded, because of the decrease in letter recall accuracy following the first letter position and the partial increase in the last position found in some experiments on non-words. Although there was not an increase in performance on the nth-letter task in the last position in all experiments, this could be explained by a strong top-down activation for existing words, which led to an error in the performance of the task. Despite of significant differences in accuracy across positions in words, we observed that the general letter recall accuracy for words was higher than for non-words and followed roughly the same trend except for minor violations, for which explanations have been offered above. s

Limitations

The results of this study should be carefully considered, due to several limitations that are present. A major source of limitation is due to the nature of the study, as it was conducted online. Online experiments are of advantage, as they are easily accessible for a large sample of people. However, as the participants could freely choose the time point of conducting the experiment, there was little control over the environment that would have been present if the experiment was conducted in the laboratory. For instance, distractors might have been present and the time of day could have negatively affected a participant's concentration. In order to prevent a lack of concentration from influencing our results, we reduced the duration of the experiment. A limitation of the short duration is however the limited data we could gather, as participants were only presented with a subset of the generated words via three experimental trials (of 40 words each). Therefore, we created word lists of comparable frequencies in order to reduce the (non-) words presented to each participant, resulting in six trials per condition. The online format further posed another limitation that regards the possibility of asking questions on the experiment in cases of unclarity. Despite indicating the supervisor's email address in the instructions of the experiment, it is logical to assume that few participants would make use of that option, send an email to the supervisor, and await a response, until they can conduct the experiment, in contrast to asking the supervisor if he were physically present.

Another limitation is posed by the way the non-words were created. The target letter was not kept at the same position in the non-word, in order to achieve the desired frequencies, however this weakens our design, in comparison to previous studies (e.g., Whittaker, 2019). Lastly, all participants included in this study were German native speakers, however they might have been bilingual. Thus, the non-words we created were not common words in German, however, they might not have been in different languages that the participant possible spoke.

Directions for Future Research

Understanding the underlying biological mechanisms of reading is of great relevance to the field of cognitive psychology, as so far, no model is present that solves the binding problem. The results of this study can be built on by expanding this experiment that further provide us with insights into the mechanisms of reading. For instance, the conceptual model proposes a top-down, left to right process of reading, however some languages are read from right to left, like Arabic or from top to bottom, like Chinese. For future studies, it might be of interest to test the conceptual network on these languages to assess the generalizability of the model to other languages, and perhaps additionally to longer letter sequences. Another interesting point to assess is how a high tail of non-words could elicit higher performance on the nth-letter task in comparison to the low tail non-words.

Additionally, as mentioned earlier, an experienced reader might have an advantage in performing on the nth-letter task. In order to control for this effect, an idea is to assess a participant's expertise in reading by admitting participants with only a certain degree to the experiment, e.g., a bachelor's degree. This way, we achieve more homogeneity in the participants' reading expertise. Furthermore, in order to avoid the negative consequences of online studies it is advisable to conduct the experiment in person. By doing so, one can ensure more environmental control that in turn increases the reliability of the results.

The complexity of the underlying biological mechanisms of reading remains a challenging topic, however our findings suggest that the conceptual network's concept of serial binding is a viable solution to the binding problem.

References

Bhouri, D. (2018). Testing the Effect of Flankers on Binding in Word and Letter Recognition (Bachelor Thesis), University of Groningen.

Buijsman, L. (2019). The Effect of Letter Position on Letter Identification Accuracy in Neural Binding (Bachelor Thesis), University of Groningen.

CELEX Centre for Lexical Information. (2001). Webcelex. Retrieved from http://celex.mpi.nl/

Davis, C. J., & Bowers, J. S. (2006). Contrasting five different theories of letter position coding: evidence from orthographic similarity effects. Journal of Experimental Psychology. Human Perception and Performance, 32(3), 535–57.

de Vries, P. H. (2016). Neural binding in letter- and word-recognition. In K. E. Twomey, A. Smith, G. Westermann & P. Monaghan (Eds.), *Neurocomputational models of cognitive development and processing: Proceedings of the 14th neural computation and psychology workshop* (pp. 17- 33). New Jersey: World Scientific, 2016.

Donelan, C. (2018). The Influence of Attentional Cues on Letter-Position Binding During Word-Recognition (Bachelor Thesis), University of Groningen.

Freericks, A. (2018). Attention and Binding in Word and Letter Recognition (Bachelor Thesis), University of Groningen.

Grainger, J., & Whitney, C. (2004). Does the huamn mnid raed wrods as a wlohe? Trends in Cognitive Sciences, 8(2), 58-59.

Hebb, D. O. (1949). *The organisation of behaviour: a neuropsychological theory*. New York: Science Editions.

Malea, M. (in preparation). Does Head Frequency Affect the Process of Letter Recall of Non-words based on the Conceptual Network Model? (Bachelor Thesis), University of Groningen.

Mathôt, S., Schreij, D., & Theeuwes, J. (2012). OpenSesame: An open-source, graphical experiment builder for the social sciences. *Behavior Research Methods*, *44*(2), 314-324. doi:10.3758/s13428-011-0168-7

McClelland, J. L., & Rumelhart, D. E. (1981). An interactive activation model of context effects in letter perception: I. An account of basic findings. Psychological Review, 88(5), 375. https://doi.org/10.1037/0033-295X.88.5.375

Mudogo, D. (2019). Binding in Letter and Word Recognition within a Conceptual Network (Bachelor Thesis), University of Groningen.

Lacroix, G. L., & Giguère, G. (2006). Formatting data files for repeated-measures analyses in SPSS: Using the Aggregate and Restructure procedures. *Tutorials in Quantitative Methods for Psychology*, *2*(1), 20-25.

Pink, D. (2019). How Veridical Word Perception Facilitates Position-specific Letter Report: A Serial Binding Account (Bachelor Thesis), University of Groningen.

Prolific. Released 2014. Prolific, Version of May 2022. London, UK. Available at: https://www.prolific.co

Qualtrics. Released 2005. Qualtrics, Version of May 2022. Provo, Utah, USA. Available at: https://www.qualtrics.com

Reicher, G. M. (1969). Perceptual recognition as a function of meaningfulness of stimulus material. *Journal of experimental psychology*, *81*(2), 275. https://doi.org/10.1037/h0027768

Schoell, L. (in preparation). The Impact of Tail Frequencies on Position-specific Letter Recalling (Bachelor Thesis), University of Groningen.

Seibel, C.M. (2018). The Role of Attention in Word and Letter Recognition (Bachelor Thesis), University of Groningen.

Schwartzkopf, R. (2019). What Position-Specific Effects Exist in the Identification of Letters in Words and in Nonwords? (Bachelor Thesis), University of Groningen.

Van der Wal, Y. (in preparation). De invloed van non-woorden op het geheugen voor letters (Bachelor Thesis), University of Groningen.

Von der Malsburg, C. (1999). The what and why of binding: the modeler's perspective. *Neuron*, *24*(1), 95-104.

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, 221–243.

Whittaker, A. (2019). At a Glance – The Role of Word Perception and Letter Position in Letter Recognition (Bachelor Thesis), University of Groningen.

Wikipedia contributors. (2022, 22 April). Pareidolia. Wikipedia. Retrieved on 25 April 2022, from https://en.wikipedia.org/wiki/Pareidolia

Yu, R. (in preparation). The Effect of Tail Frequencies on Binding of Specific Position in (Non)Word and Letter Recognition (Bachelor Thesis), University of Groningen.