The optimal use of non-optimal letter information in foveal and parafoveal word recognition

Jean-Baptiste Bernard, Eric Castet

The optimal use of non-optimal letter information in foveal and parafoveal word recognition

01/01/2019

Jean-Baptiste Bernard[1,2,3] and Eric Castet[1,2]

[1]Laboratoire de Psychologie Cognitive, Centre National de la Recherche Scientifique, UMR 7290

[2]Aix Marseille Université

[3]Brain and Language Research Institute

Corresponding author:

Jean-Baptiste Bernard, PhD

Laboratoire de Psychologie Cognitive

Bâtiment 9 Case D, 3, place Victor Hugo.

13331 Marseille Cedex 3

Email: bernard_jb@hotmail.com

Abstract

Letters and words across the visual field can be difficult to identify due to limiting visual factors such as acuity, crowding and position uncertainty. Here, we show that when human readers identify words presented at foveal and para-foveal locations, they act like theoretical observers making optimal use of letter identity and letter position information independently extracted from each letter after an unavoidable and non-optimal letter recognition guess. The novelty of our approach is that we carefully considered foveal and parafoveal letter identity and position uncertainties by measuring crowded letter recognition performance in five subjects without any word context influence. Based on these behavioral measures, lexical access was simulated for each subject by an observer making optimal use of each subject's uncertainties. This free-parameter model was able to predict individual behavioral recognition rates of words presented at different positions across the visual field. Importantly, the model was also able to predict individual mislocation and identity letter errors made during behavioral word recognition. These results reinforce the view that human readers recognize foveal and parafoveal words by parts (the word letters) in a first stage, independently of word context. They also suggest a second step where letter identity and position uncertainties are generated based on letter first guesses and positions. During the third lexical access stage, identity and position uncertainties from each letter look remarkably combined together through an optimal word recognition decision process.

INTRODUCTION

It only takes tens of milliseconds for humans to identify words, a combination of letters arranged in a given order to form the essential language unit. Several decades of research suggested that visual word recognition is preceded by a letter recognition step, which is itself preceded by the extraction of letter visual features (see Carreiras, Armstrong, Perea, & Frost, 2014 for a recent review). Identification of word sub-units such as syllables, graphemes or phonemes can also take place between letter and word recognition (Balota, Yap, & Cortese, 2006). Given this cascade of processes, an alteration of letter recognition performance can be critical for word recognition performance. This is especially true for words presented outside the fovea where letter and thus word recognition performance are strongly altered (Latham & Whitaker, 1996). Even in foveal word viewing recognition performance is usually impaired for many letters. This is because letter recognition performance is not homogeneous across the visual field and quickly drops with visual eccentricity (Legge, Mansfield, & Chung, 2001). However, visual limitations can be compensated by the knowledge of word lexicon or sentence context, allowing the recognition of most words within one single foveal fixation, although eccentric letters are not perfectly visible (Legge, Klitz & Tjan, 1997).

Acuity, crowding and positional uncertainty have been suggested to be the main visual factors that limit letter recognition across the visual field (He, Legge, & Yu, 2013; Legge et al., 2007; Pelli & Tillman, 2008; Yu, Legge, Wagoner, & Chung, 2014). Low acuity impairs letter identification in the para-fovea and in the periphery when a letter is presented alone or surrounded by other letters (Wertheim, 1980; Westheimer, 1979). Crowding impairs letter identification when a target letter is surrounded by other letters (Bouma, 1970; Levi, 2008; Whitney & Levi, 2011). Crowding supposedly constrains the visual system to integrate together visual features coming from a target letter and from its neighbor letters (Bernard & Chung, 2011; Pelli, Palomares, & Majaj, 2004; Pelli & Tillman, 2008). The negative impact of crowding on letter recognition is directly proportional to visual eccentricity (i.e. the Bouma law (Bouma, 1970)). It is more deleterious on reading performance than the impact of visual acuity when letters are presented at a usual print-size (Pelli et al., 2007; Yu et al., 2014). Position uncertainty is another visual constraint that decreases the ability of subjects to correctly localize single or crowded objects such as letters (Chung & Legge, 2009; Levi & Tripathy, 1996; Strasburger, 2005; Strasburger & Malania, 2013). Localization errors can be represented by a normal distribution centered on the letter target (Chung & Legge, 2009; Gomez, Ratcliff, & Perea, 2008; Levi & Tripathy, 1996). Standard deviation of this distribution varies as a function of certain parameters such as time duration or eccentricity (Chung & Legge, 2009; Gomez et al., 2008; Michel & Geisler, 2011). Crowding and localization are tightly connected as it has been shown that crowding increases positional uncertainty for letters (Harrison & Bex, 2016; van den Berg, Johnson, Martinez Anton, Schepers, & Cornelissen, 2012). Visual acuity, crowding, and positional uncertainty are thus the main visual factors which theoretically limit the identification of letters within a word.

Therefore, a complete word recognition model should represent (1) the detection and integration of visual features to build letters, and (2) the integration of letter-level information to reach word identification. This is the basic theory that is described by the original interactive activation model (McClelland & Rumelhart, 1981) and its descendants (Coltheart, Rastle, Perry, Langdon, & Ziegler, 2001; Davis, 2010; J. Grainger & Jacobs, 1996).[1] In this theory, the first feature-to-letter step is still unclear because letter features are not yet defined despite decades of research (Fiset et al., 2008; Jonathan Grainger, Rey, & Dufau, 2008). As a consequence the recognition of crowded letters within a letter string involving the combination of target and flanker features (Pelli et al., 2004) is even harder to define. Thus this step is usually skipped or grossly represented in implemented word recognition models. The letter-to-word process through a lexical access has been largely studied in behavioral (essentially through priming experiments) and in brain activation studies. Lexical access is usually represented by a competition across time between words of the lexicon that are perceptually similar: When a reader starts to identify a word, visual information about its letters increases with

[1] Note that these models do not separate letter and word identification steps although (Pelli, Farell, & Moore, 2003) and our results suggest otherwise.

time, consequently increasing the activation of the word at the expense of its orthographic neighbors. This competition is another common basis of word recognition models (Norris, 2013), with recent models even describing the accumulation of letter information across time and its use to identify words (Adelman, 2011; Norris, 2006; Norris & Kinoshita, 2012; Norris, Kinoshita, & van Casteren, 2010). It is clear that two categories of letter-level visual information influence the discrimination between a word and its orthographic neighbors: identity and position information (Davis, 2010; Gomez et al., 2008; Norris, 2006; Norris et al., 2010). Behavioral results on visual crowding described in the previous paragraph show that under crowded conditions, human observers can usually extract a limited amount of letter identity and position information, this information depending on many visual factors such as presentation duration, eccentricity or spatial configuration (Chung & Legge, 2009; Legge et al., 2001). Based on the strong influence of visual crowding on letter recognition it is rather surprising that the influence of crowding on letter recognition is almost never (or very superficially) taken into account in current implemented models (Norris, 2013). More generally we argue that the influence of visual crowding on letter recognition cannot be bypassed if we want to predict errors that occur during the recognition of long words (length > 5 letters). This is even more critical for words presented in the parafovea or in the periphery as the deleterious effect of visual crowding increases with visual eccentricity.

Bayesian models have been able to predict human perception and oculomotor performance and behaviors in vision science (Geisler, 1989; Geisler, 2011; Kersten, Mamassian, & Yuille, 2004; Legge et al., 1997; Najemnik & Geisler, 2005; Renninger, Verghese, & Coughlan, 2007; Weiss, Simoncelli, & Adelson, 2002). Theoretically, Bayesian observers act like ideal observers who make decisions based on an optimal use of the incomplete available visual information (given the limitations of the visual system). Concerning word recognition, Bayesian theories have been implemented and tested to predict reading performance in foveal and parafoveal reading either with (Bicknell & Levy, 2010, 2010; Legge et al., 2001; Legge, Hooven, Klitz, Stephen Mansfield, & Tjan, 2002) or without (Legge et al., 2001; Pelli et al., 2003) eye movements. These models defined identity uncertainty as the only visual factor limiting foveal and parafoveal letter recognition. Legge et al (Legge et al., 2001) calculated the performance of a Bayesian lexical matching algorithm and found that the model was sufficient to account for reading performance in foveal vision. The "Bayesian reader" (Norris, 2006; Norris et al., 2010) used only letter identity uncertainty in its first version (Norris, 2006). Later both letter identity and letter position uncertainties were used to simulate foveal word recognition processes (Norris et al., 2010). The "Bayesian Reader" describes the accumulation and optimal integration of letter feature information, making letter identity and letter position (and thus word identity) less and less uncertain over time. The Bayesian Reader predicts the (logarithmic) effect of word frequency on foveal word reaction times as well as some typical priming effects that modulate lexical decision performance. In the implemented version of this model all letters in a word are assumed to be equally visible, an acceptable hypothesis for the foveal recognition of short words.

In the present study, as in previous Bayesian word recognition models, we assumed that human readers make an optimal use of available letter information during the lexical access (Dennis Norris, 2006). However, we also took into account that an automatic non optimal letter-processing stage (occurring independently for all letters) precedes lexical access (Pelli et al., 2003). The model that we implemented and tested here is based on these assumptions. It can be broken down in three stages: The first stage is the automatic non-optimal letter-processing stage proposed by Pelli et al. (2003). The two following stages are optimal. Our second stage consists in the extraction of letter position and identity uncertainties based on the first-step letter “recognition” process. Our third stage corresponds to the lexical access and is the integration position and identity uncertainties from each letter in order to identify the presented word. We tested our model with word recognition tasks in parafoveal conditions because they are ideal to create strong perceptual uncertainties with long or unlimited durations without inducing artificial visual noise as in foveal viewing. We can thus judge the validity of our model by directly using word recognition performance. This also allows us to generalize the validity of Pelli et al’s theory in peripheral viewing conditions and to understand the mechanisms of word recognition for subjects who cannot use their central vision.

The important novelty of our approach is the way we tested our model: We first behaviorally quantified letter identity and position uncertainties caused by acuity, crowding and position uncertainty. Currently implemented word recognition models usually assume that letter identity uncertainty is homogeneous within the presented word, which is a bold approximation given the large slope of the relationship between eccentricity and crowded letter recognition (Legge et al., 2001). Here we bypassed this strong limitation by estimating position and identity letter uncertainties for different letter positions in five individual subjects (Experiment 1). These psychophysical measurements were made in the absence of any linguistic contextual constraints. They allowed us to implement an ideal-observer model that simulated the optimal use of both letter identity and position information for each observer following a non-optimal letter guess (Pelli et al., 2003). These predictions were compared with word recognition data from the same five subjects (Experiment 2). To anticipate our results, we showed that word recognition errors and corresponding letter errors made by our ideal-observer model correspond to errors made by human readers. Our work suggests that word recognition can be described as the succession of three stages: A first stage where expert readers automatically try to identify all word letters (Pelli et al., 2003) thus leading to "letter guesses", a second stage where letter-identity and letter-position information within a word are determined for each letter, and a third lexical access stage where both types of letter information are optimally combined together.

Based on the successful comparison between our ideal-observer model and psychophysical data, our work confirm the plausibility of such an optimal word recognition mechanism (stages 2 and 3) following the automatic and non-optimal word letter recognition step (stage 1).

METHODS

Subjects

Five subjects (age: 22-37) participated in Experiment 1 (letter recognition experiment) and the same five subjects participated in Experiment 2 (word recognition experiment). Subjects were all native-born French speakers. The research followed the tenets of the Declaration of Helsinki and was approved by the Ethical Committee for Protection of Human Subjects at the Aix-Marseille Université. Written informed consent was obtained from each observer after the nature and purpose of the experiment had been explained.

Apparatus

Stimuli were displayed on a 21-inch CRT color monitor (ViewSonic P227f, refresh rate = 120 Hz, resolution = 1152 x 854 pixels). A PC computer running custom software developed in Python with the Psychopy library (Peirce, 2007) was controlling the display. Observers sat in a comfortable chair at a viewing distance of 40 cm (screen visual angle: 50.8° x 37.7°) with a forehead rest to stabilize their position. An Eyelink 1000 Tower Mount eyetracker (SR Research Ltd., Mississauga, Ont., Canada) was also connected to our system to control observers' gaze position in Experiments 1 (letter recognition) and 2 (word recognition). In both experiments letters and words were displayed in black (luminance: 0.3 cd/m2) on a light gray background (luminance: 60 cd/m2).

Letter strings and words used in experiments

Letter strings used in Experiment 1a (3-letter strings) and 1b (5-letter strings) were made of letters randomly chosen among the 26 letters of the alphabet. Word lemmas used in Experiment 2 were randomly extracted from three sets of 500 lemmas (500 5-letter words, 500 7-letter words and 500 9-letter words). Each word had a lexical frequency larger than 15 occurrences per million of words based on the Lexique3 corpus (New, Pallier, Brysbaert, & Ferrand, 2004) to avoid the presence of words which could be unknown to the subjects. Importantly, all French trigrams and French words were chosen without any accent letters.

Experimental protocol

Each subject ran three successive experiments that measured foveal and parafoveal letter recognition (Experiments 1a and 1b) and word recognition (Experiment 2) performance. Letters and words were presented at different positions on an invisible horizontal lined centered on the middle of the screen. Experiments 1a and 1b were run to measure letter recognition performance for letters presented within trigrams (Experiment 1a) and within pentagrams (Experiment 1b). Experiment 2 was run to measure word recognition performance. In each experiment and for every subject the x-height letter print-size (Legge & Bigelow, 2011) was fixed at 1.38°. We assume that this print-size is large enough to avoid the effect of acuity on letter recognition in our experiment as Yu et al (Yu et al., 2014) found a very small acuity effect on letter recognition with a print-size of 0.55° and a protocol similar to ours. The non-proportional Courier font was used to display letters and words so that the letter slot positions in Experiments 1a and 1b corresponded to the same letter slot positions as in Experiment 2. Presentation duration was fixed at 250 ms, the approximate average duration for a reading fixation (Rayner, 1998)

(1) Experiment 1: Letter recognition

Experiment 1 was similar to experiments used to measure what previous vision researchers called visual span profiles (Legge et al., 2001; Legge et al., 2007): the recognition performance of letter trigrams (three adjacent letters) presented at different retinal locations. Full report of letter trigrams is useful because it allows the measurement of error rates for interior letters (i.e. letters with one flanker on the left and one flanker on the right) and for exterior letters (i.e. letters with only one flanker on the left or only one flanker on the right). Observers were required to report verbally the three letters from left to right: Therefore we called left, center, and right the three possible letter positions[2]. **Figure 1a** and **Figure 1b** describe the temporal course of Experiments 1a and 1b: Observers were asked to fixate between two dots centered on the middle of the screen (distance between the two dots: 3°). Gaze location was measured to control for steady fixation between the two dots (tolerance: ±0.5°). In Experiment 1a, when the observer was ready for the trial, he/she pressed the button of a hand-held joypad. This started the trial: a letter trigram (a string of three random letters with standard letter-to-letter spacing) was centered at one of 13 possible positions across the horizontal meridian. Before the trigram display, three hashtag symbols were displayed at the future letter positions until the subject pressed the button. A backward mask with the same three hashtag symbols immediately followed the trigram presentation. Each observer was required to report the three presented letters verbally from left to right as the experimenter recorded the report. Each subject ran one session (approximately 1h) which consisted of 4 blocks of 65 trials (13 locations x 5 repetitions). For each subject, a total number of 20 trigram trials was run for each location condition.

The experimental protocol of Experiment 1b was similar to the protocol of Experiment 1a (recognition of letter trigrams) and is described in **Figure 1b**. The main difference was that a string of 5 letters was displayed on each trial. Subjects had to report the three interior letters (i.e. the interior trigram) of the pentagram. These three interior letters could be displayed at one of the 13 positions that had been used in Experiment 1a. They were also called the left, central, and right letters (i.e. with respect to the central letter) of the interior trigram . Five hashtag symbols were first displayed to indicate the future letter positions and a backward mask was presented comprised of the same hashtag symbols. As in Experiment 1a each observer's answer was recorded. Each subject ran one session (approximately 1h that consisted of 4 blocks of 65 trials (13 locations x 5 repetitions). For each subject a total number of 20 pentagram trials was run for each (location x duration) condition.

[2] Note that this categorisation is not done for the classical visual span profile measurement.

In both experiments 1a and 1b, location of the letters to the left and to the right of the central letter ("x" and "z" in figure 3) is referred to hereafter as either the relative position within the interior trigram or as the interior position for short.

(2) Experiment 2: Word recognition

Figure 1c describes the temporal course of Experiment 2: Five, seven, or nine-letter words were briefly presented, centered on different positions (five possible center positions: the fixation locus, two letter slots on the left, two letter slots on the right, four letter slots on the left, and four letter slots on the right). Presentation duration was 250 ms (as in Experiment 1). The future position of a word was indicated in advance by several hashtag symbols (5, 7 or 9 symbols) until the observer pressed the button and the word was displayed. Right after the word display a backward mask made of the same number of hashtag symbols replaced the word. The observer's verbal report was stored by the experimenter. Each subject ran two sessions (approximately 1h per session). Each session was made of 6 blocks. Each block corresponded to the display of words of a given word length (i.e. the word length was blocked) and was made of 25 words (5 positions x 5 repetitions). Subjects were not allowed to give non-word responses. Before the beginning of a block subjects were informed of the size of the words. If the subject was reporting a word with a different length the experimenter would collect this response anyway. For each subject 20 words were displayed for each (position x word length) condition.

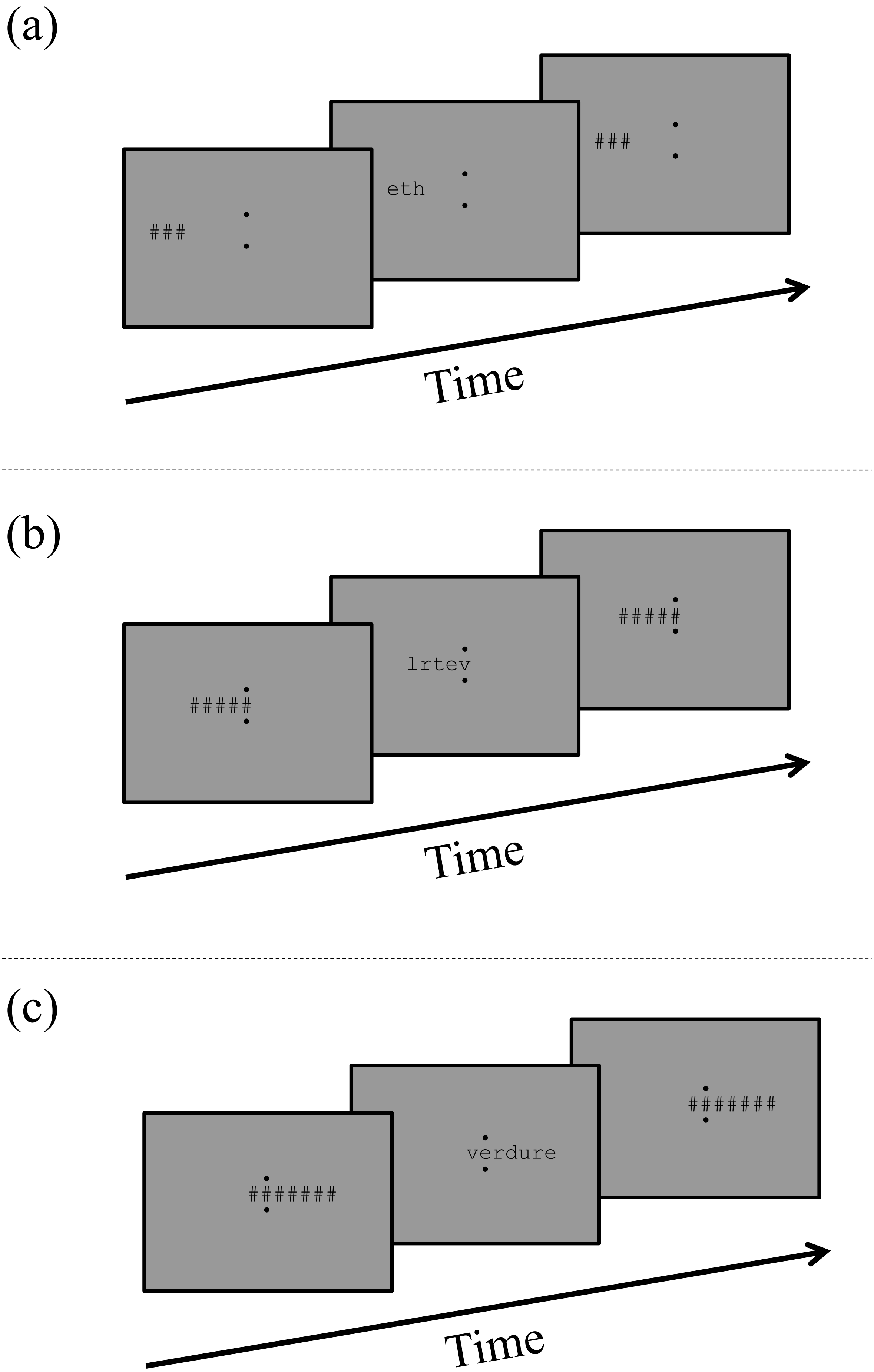

(a)
###
eth
###
Time
(b)
#####
lrtev
#####
Time
(c)
#######
verdure
#######
Time

Figure 1: Experimental protocol for Experiments 1a, 1b, and 2.

Observers were asked to fixate between the two central dots before (a) a letter trigram (Experiment 1a), (b) a letter pentagram (Experiment 1b), or (c) a 5-letter, 7-letter, or 9-letter French word (Experiment 2) was briefly displayed for 250 ms as soon as they pressed a button. At the end of the trial, observers reported the letter string/word to the experimenter (the 3 interior letters when a pentagram was displayed, Experiment 1b).

RESULTS

Letter recognition profiles

Figure 2 shows an example of a letter recognition profile for one of our five subjects. The average recognition performance for each dot is based on twenty trials. Letter recognition performance is highest for letters presented at the fovea and decreases with eccentricity. Legge et al (Legge et al., 2001) showed that two half-Gaussian functions can efficiently quantify the effects of eccentricity on letter recognition rates. This is particularly useful because it improves the quality of the collected data and it uses only 3 parameters to define a visual span profile. In consequence, each (visual hemifield h, trigram interior position i) association can be characterized by a *letter recognition profile* function defined as a function f of horizontal eccentricity:

$letter\_recognition\_rate = f_{(h,i)}(ecc)$, f being a half-Gaussian function with $f_{(h=L,i)}(0) = f_{(h=R,i)}(0) \leq 1$ and $\lim_{ecc\to\infty} f_{(h=L,i)}(ecc) = \lim_{ecc\to\infty} f_{(h=R,i)}(ecc) = {}^{1}/_{26}$. These equations mean that, for a given interior position, the foveal value for the left visual field corresponds to the foveal value for the right visual field. For the left and the right visual field, the chance level for infinite eccentricity is ${}^{1}/_{26}$. These mathematical rules were used to fit the raw data into letter recognition profiles for each subject and each relative position within the letter trigram (Experiment 1a) or the three central letters of the pentagram (Experiment 1b). Profiles for each subject, each relative position within the trigram or the three central letters, and each experiment are shown in **Figure 3**. They show a much better identification performance for trigram letters compared to pentagram letters (37% vs. 65% on average) because of the increase of neighbour letters in the pentagram task.

Based on our data average points we investigated the effects of different factors on letter recognition performance. Two generalized linear mixed-effect models for binary responses (function glmer of the lme4 package in the R language and environment (R Development Core Team, 2013)) were performed to analyse letter recognition performance in Experiments 1a and 1b. The random effect was the subject factor and the fixed effects were the visual hemifield side, the eccentricity, and the relative position within the trigram made by the three central letters of the pentagram. There were two interaction terms: visual hemifield* relative position within the interior trigram and eccentricity*relative position within the interior trigram. The dependent variable was the letter recognition error variable (0 or 1). Results of the analysis are shown in Table 1a for Experiment 1a and Table 1b for Experiment 1b.

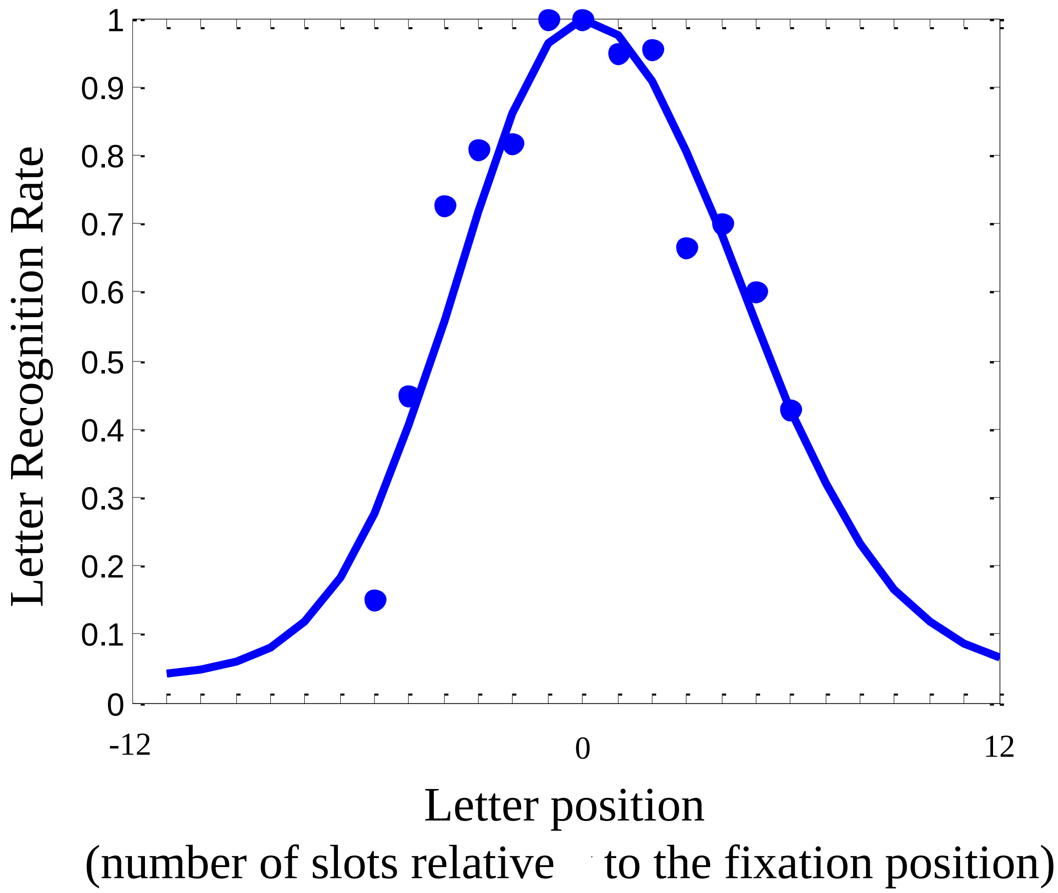


Figure 2: Example of letter recognition profile

Each dot represents the average recognition rate for each letter position (from 20 trials). The curve represents the two connected half-Gaussian functions that offer the best fit for the given dots in the left and in the right visual field.

Our analyses confirm what can be observed in **Figure 3**: There are differences in letter recognition performance based on eccentricity, relative position within the trigram made by the three central letters (3 coloured letters in figure 3), and visual hemifield as already described in different studies dealing with peripheral recognition of crowded letters (Legge et al., 2001). Confirming this observation our two mixed effect model analyses show significant effects of visual hemifield, internal position, and visual eccentricity on crowded letter recognition rates. In both experiments we found a significant advantage of the right visual hemifield over the left visual hemifield (Bouma, 1973; Bouma & Legein, 1977; Nazir, Heller, & Sussmann, 1992): + 7% on average for the right visual field. We also found a significant decrease of letter recognition performance with visual eccentricity (Bouma, 1973): - 18% on average per letter slot. The significant difference between left, central, and right interior positions is different in both experiments: In Experiment 1a (trigram presentation) there is a disadvantage for the central letter compared to the left and right letters because of the lack of crowding for the two outer letters (Bouma, 1973; Legge et al., 2001). In Experiment 1b (pentagram presentation), we found a significant left-to-right gradient effect: Recognition performance is significantly better for the left letter compared to the center letter, and for the center letter compared to the right letter. This small effect could represent a leftward bias (sequential report) already reported in similar letter string experimental measurements (Whitney, 2008). We also found a significant interaction between the relative position within the trigram and the hemifield in Experiment 1a: The outer letter (left letter in the left visual field and right letter in the right visual field) is easier to identify, an effect already shown in previous study (G. E. Legge et al., 2001). This effect is only significant for the right letter of the trigram in Experiment 1b. Finally, a significant interaction between the eccentricity and the relative position within the trigram was found in Experiment 1a: The effect of visual eccentricity is larger for the left than for the right letter.

These results confirm that crowded letter recognition performance is dependent on three key parameters that have been previously described: visual hemifield, interior position, and eccentricity.

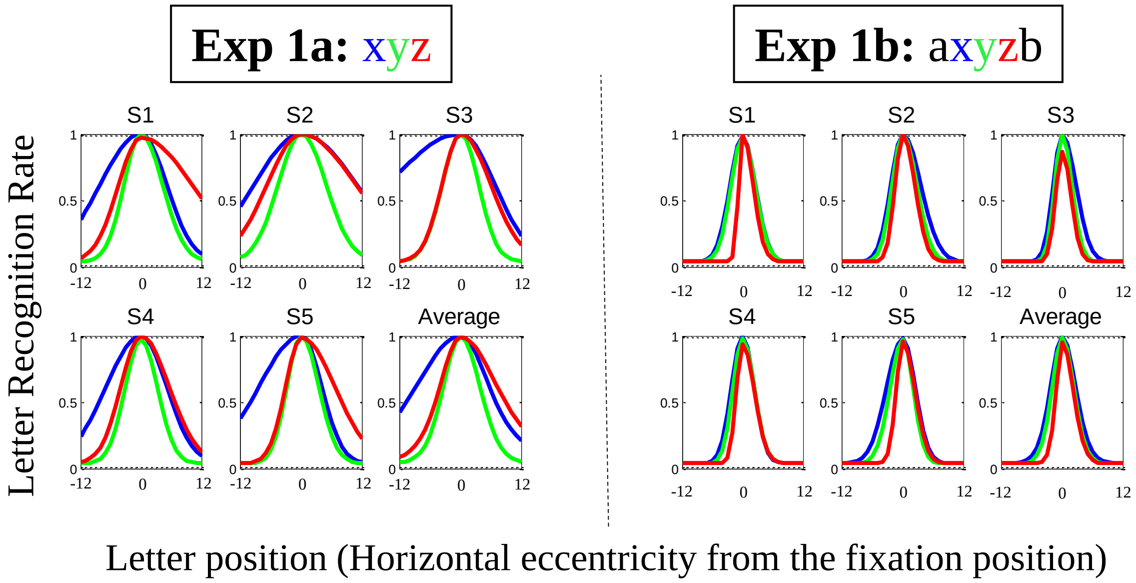


Figure 3 : Letter recognition profiles. The left and right boxes represent the Letter recognition profiles for the two different letter recognition experiments: Experiment 1a and Experiment 1b. Each box is made of six smaller boxes (one for each subject and one for the average data). Each curve, made of the two connected half-Gaussian functions, represents the subject (or average) letter recognition profiles for the three different relative interior positions (left position in blue ("x"), central position in green ("y"), and right position in red ("z")) within each trigram (exp 1a) or pentagram (exp 1b).

Table 1a. Exp 1a

Fixed effects	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	0.6210283	0.2404979	2.582	0.009816
visual hemifield (right)	0.3500869	0.0808396	4.331	1.49e-05
Relative position within the trigram (left)	2.0014702	0.1092974	18.312	<2E-16
Relative position within the trigram (right)	0.2867366	0.0855667	3.351	0.000805
eccentricity	-0.7357411	0.0269314	-27.319	<2E-16
visual hemifield (right):Relative position within the trigram (left)	-1.3902488	0.1277637	-10.881	<2E-16
visual hemifield (right):Relative position within the trigram (right)	1.0746496	0.1222340	8.792	<2E-16
Relative position within the trigram (left):eccentricity	0.1786149	0.0413008	4.325	1.53e-05
Relative position within the trigram (right):eccentricity	0.1132925	0.0392498	2.886	0.003896

Table 1b. Exp 1b

Fixed effects	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	-0.5293753	0.1230699	-4.301	1.70e-05
visual hemifield (right)	0.2767427	0.0852324	3.247	0.00117
Relative position within the trigram (left)	0.3937151	0.0843066	4.670	3.01e-06
Relative position within the trigram (right)	-1.2238469	0.0938575	-13.039	<2E-16
eccentricity	-0.9063045	0.0303009	-29.910	<2E-16
visual hemifield (right):Relative position within the trigram (left)	-0.1061905	0.1184965	-0.896	0.37017
visual hemifield (right):Relative position within the trigram (right)	1.0813470	0.1263632	8.557	<2E-16
Relative position within the trigram (left):eccentricity	0.0657647	0.0409290	1.607	0.10810
Relative position within the trigram (right):eccentricity	0.0600806	0.0438134	1.371	0.17029

Table 1: Fixed effects results of the generalized linear mixed-effects models to predict letter recognition rates in Experiment 1a (Table 1a) and Experiment 1b (Table 1b). In these models, the reference values were: Left visual hemifield, central "interior position" within the trigram made by the 3 central letters (coloured letters in figure 3), and a visual eccentricity of 3 letter slots. No hemifield was assigned to the letters on the midline.

Separating letter identity and letter position errors

Previous studies have shown that crowded letter recognition errors can be divided in two types of errors (Strasburger & Malania, 2013; Wang, He, & Legge, 2014; Yu et al., 2014) : location errors (a neighbor letter is reported instead of the target letter) and identity errors (an error that is not a location error). To separate both types of letter errors we first assumed that identity and mislocation errors are independent which is a plausible hypothesis (Hanus & Vul, 2013) despite some controversies (Bernard & Chung, 2011; Freeman, Chakravarthi, & Pelli, 2012). Second we assumed that letter position, for a given spatial horizontal position, can be represented by a Gaussian probability density distribution centered on the real position of the letter (Chung & Legge, 2009; Levi & Tripathy, 1996). Interestingly the standard deviation of this position noise has been shown to increase almost linearly with visual eccentricity (Chung & Legge, 2009; Levi & Tripathy, 1996; Michel & Geisler, 2011). Therefore a single mislocation coefficient α (a proportionality coefficient if we consider positional uncertainty as null at the fovea) is sufficient to characterize the position uncertainty of letters at any horizontal position across the visual field. **Figure 4** shows how we determined this coefficient ($\propto$) from each letter recognition profile: First we removed mislocation errors from identification errors among letter errors following what has been done in previous studies (He et al., 2013; Wang et al., 2014; Yu et al., 2014). Letter mislocation for a given letter occurs when the letter is correctly identified but not at the true letter slot position (For instance if the trigram abc was presented, the letter b was considered as mislocated if reported at the first or last slot by the subject but not at the center slot). **Figure 4a** shows the new letter recognition profiles (in red) without mislocation errors. This new profile is critical to determine letter identity uncertainty. We call it the *letter identity recognition profile.* It is logically higher than the letter recognition profile with mislocation errors (in blue) because it does not consider letter mislocations as letter recognition errors. Once these profiles were extracted for each subject we determined the optimal *mislocation coefficient* (α) (see **Figure 4b** and **Figure 4c**) that offers the best fit (dashed blue line in **Figure 4a**) to predict the *letter recognition profile (in blue) from the letter identity recognition profile (in red).* Note that this prediction offers a very high correlation coefficient (r^2>0.90) for each obtained mislocation coefficient (α). The mathematical method to calculate a letter recognition profile from an identity letter recognition profile and a mislocation coefficient (α) is detailed in Appendix 1.

Letter mislocations were determined for the three relative positions in Experiment 1a and only for the central position in Experiment 1b. Indeed in Experiment 1b (pentagram presentation), subjects did not report all five letters, and it was impossible to know if the subject correctly identified a letter at a wrong position because he/she was naming only 3 letters. For instance, a subject could report a mislocalized letter displayed at the first or last position of the pentagram but we could not know if it was the case because the subject was only naming the three central letters. However, we assumed that for the letter reported at the central slot, most mislocations were directed only towards the 2nd and 4th letters and that we could ignore other non-reported letters. The values of the different mislocation coefficients (α) are given in Appendix 3. They characterize letter position uncertainties for the different subjects and the different conditions. The three coefficients that characterize identity letter recognition profiles (The standard deviation of the left half Gaussian distribution, the standard deviation of the right half Gaussian distribution, and the amplitude of both Gaussian distributions) for Experiments 1a and 1b are given in Appendix 4. They characterize letter identity uncertainties for the different subjects and the different conditions.

In sum, *letter identity recognition profiles* and *mislocation coefficients* characterize the uncertainty about letter identity and position information extracted for every subject in one fixation of 250 milliseconds. For each subject this information is available for different visual eccentricities and internal positions. Experiment 1a gives parameter values for exterior letters (only one letter flanker on one side) and Experiment 1b gives parameter values for interior letters flanked with at least two letters on each letter side. To simplify, letter identity recognition profiles will be called letter identity profiles in the following text.

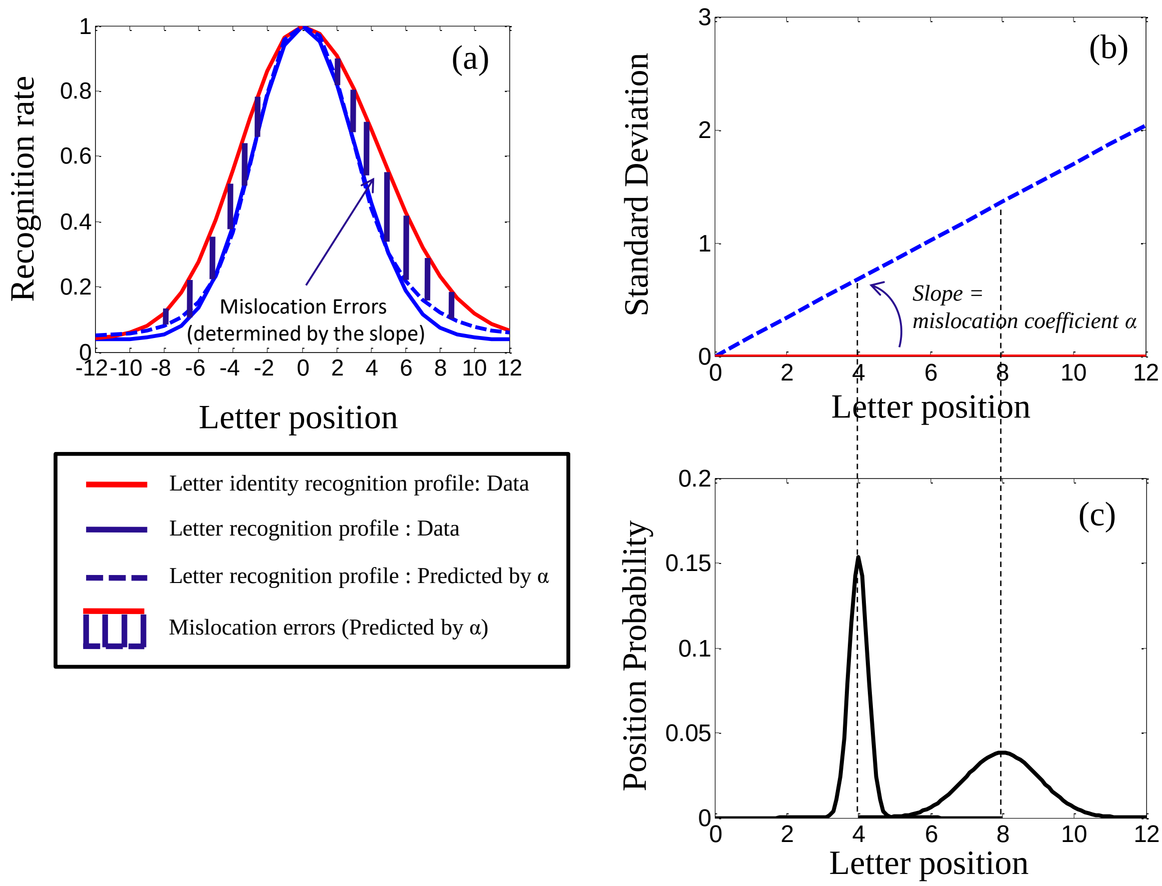


Figure 4: Determination of the mislocation coefficient α

Figure 4a shows an example of a letter recognition profile (blue curve) for a given subject and a given relative position. We first removed mislocation errors from identification errors to obtain the letter identity recognition profile (red curve). In consequence, the number of mislocation errors is represented by the difference between both curves. The mislocation coefficient α is the slope (shown in Figure 4b) of the regression determining the standard deviation of the Gaussian localization distribution as a function of the visual eccentricity (as shown by 2 eccentricity examples in Figure 4c). The optimal mislocation coefficient is the one predicting the best estimate of the increase of mislocation errors, here the dashed area. (In a perfect case, the predictive blue dashed curve would perfectly correspond to the blue solid curve).

Experiment 2: Word recognition

Figure 5a shows the word recognition performance (words made of 5, 7, and 9 letters) calculated for each subject at different eccentricities (20 word presentations for each data point). Mean performance across subjects is also indicated in the figure. Similarly to letter recognition performance, word recognition performance clearly decreases with eccentricity. A generalized linear mixed-effect model for binary responses (function glmer of the lme4 package in the language and environment R (R Development Core Team, 2013)) was performed to investigate the effects of relevant factors (hemifield, eccentricity, and word length) on word recognition performance. The random effect was the subject factor. The dependent variable was the word recognition error variable (0 or 1). Results of the analysis are shown in **Erreur ! Source du renvoi introuvable.**. As previously shown in the literature word recognition performance significantly decreases when horizontal eccentricity increases and when word length increases (Brysbaert, Vitu, & Schroyens, 1996). Words were also easier to identify when they were presented in the right visual field, with an average optimal position located between 0 and 2 slots on the left of the center of the word. This position is usually called the Optimal Viewing Position (Brysbaert et al., 1996; O'Reagan & Jacobs, 1992).

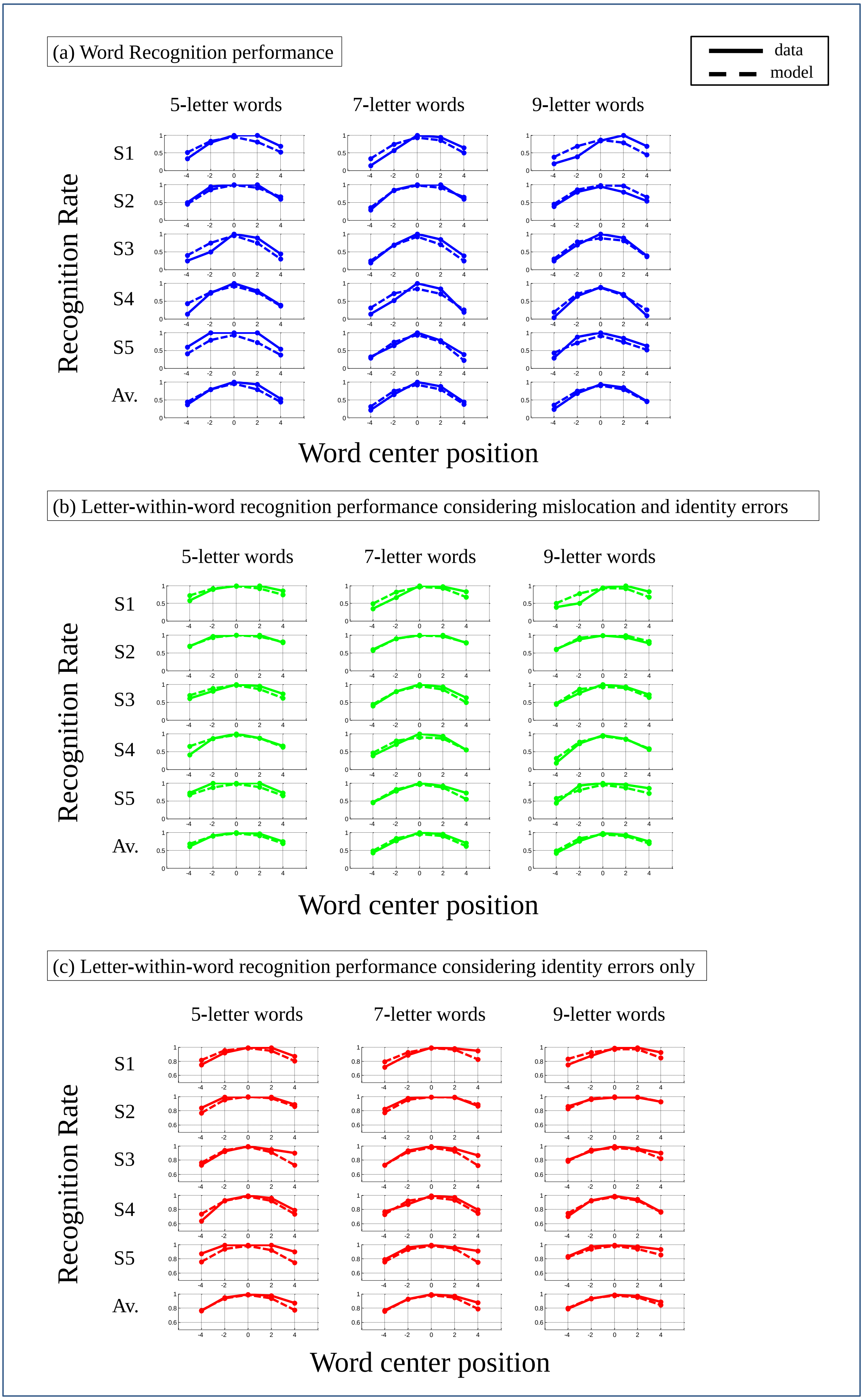
Exp 2				
Fixed effects	**Estimate**	**Std. Error**	**z value**	**Pr(>|z|)**
(Intercept)	4.360246	0.218799	19.928	2.00E-16
Visual hemifield (right)	1.111439	0.0746	-14.899	2.00E-16
Word length	-0.132066	0.021718	-6.081	1.19E-09
eccentricity	-1.0693	0.038412	-27.838	2.00E-16

Table 2: Fixed effects results of the generalized mixed-effects model to predict word recognition rates in Experiment 2. The reference values were: Left visual field and an eccentricity of 3 letter slots.

Letter errors

During Experiment 2, subjects' responses were entered letter-by-letter by the experimenter in the experimental program. This allowed us to describe the patterns of letter errors made by the subjects during the word identification task and to calculate the proportion of mislocalized and incorrectly identified letters. **Figure 5b** shows the proportion of letters that were correctly reported at their actual slot position. The letter errors that are represented here are *identity and position errors*: In this case a mislocalized letter is considered as an error. **Figure 5c** shows the proportion of letters that were correctly reported but this time at any letter position within the word. These errors approximate the letter identity errors only: In this case a mislocalized letter is not considered as an error. The goal of the word recognition model presented below is to account for the proportion of word recognition errors at different visual eccentricities, as well as for the different ratios of identity and mislocation letter errors.[3]

[3] In cases where the participant's response had a different number of letters than the presented word, the slot position was defined relatively to the first letter of the word. For instance, if the word “timer” was presented while the word “trimer” was reported, only one letter (“t”) was considered to be reported at a correct position.



(a) Word Recognition performance
data
model
5-letter words
7-letter words
9-letter words
Recognition Rate
S1
S2
S3
S4
S5
Av.
Word center position
(b) Letter-within-word recognition performance considering mislocation and identity errors
5-letter words
7-letter words
9-letter words
Recognition Rate
S1
S2
S3
S4
S5
Av.
Word center position
(c) Letter-within-word recognition performance considering identity errors only
5-letter words
7-letter words
9-letter words
Recognition Rate
S1
S2
S3
S4
S5
Av.
Word center position

Figure 5: Word recognition performance across the visual field

For different word lengths (5,7, and 9 letters) and different positions across the visual field (words centered on -4, -2, 0, 2, and 4 slots with respect to the fixation position), we plotted: (a) Individual average word recognition rates, (b) individual average letter recognition rates considering identity and mislocation errors (a response is counted as correct if the letter is correctly identified whatever its reported location within the word), and (c) individual average identity recognition rates considering identity errors only (a response is counted as correct only if a correct identification occurs at the actual location). In other words, a correctly identified but mislocated letter is counted as an error in (b) and as a correct response in (c). Each data point represents 20 trials. Raw data and predictions from our model are indicated.

A model to predict word recognition performance based on identity and position letter uncertainties

Comparison with other word recognition models

Our model extends an important precursor model (Legge et al., 2001) which defined an ideal word recognition observer based on letter identity uncertainty only[4]. Here we took an additional step forward by including letter position uncertainty and by testing if this ideal-observer model could explain word recognition error rates and patterns of letter recognition errors that occur during foveal and parafoveal word recognition tasks. This was inspired by some recent work on letter position uncertainty (Chung & Legge, 2009) and the possibility to code letter position within a word as an absolute letter position distribution. The overlap model (Gomez et al., 2008), the spatial coding model (Davis, 2010), or the latest versions of the Bayesian Reader (Norris & Kinoshita, 2012; Norris et al., 2010) implemented letter position uncertainty in the same way and showed that it can explain some foveal word recognition priming results (for instance "leakage" of letter identity to nearby positions) that cannot be explained with relative coding of letter position. These models assume that each letter of the word is equally perceptible (only the interior letters in the spatial coding model). This is a plausible assumption for short words presented foveally, but as shown in Figure 2, letter recognition rate quickly decreases as a function of visual eccentricity. In order to model word recognition performance for foveal and parafoveal words, we argue that it is crucial not to ignore these large discrepancies concerning letter identity uncertainties. This would be equivalent to ignoring visual crowding despite its evident influence on word recognition and reading performance (Frömer et al., 2015; Pelli et al., 2007; Risse, 2014). In consequence, a complete word recognition model needs to take the visual limitations of crowding on letter identification into account. The accurate representation of identity and position uncertainty of crowded letters is the key difference between our model and other models of word recognition.

[4] This model also suggests that letter recognition uncertainty is based on a prior letter recognition guess.

The current version of our model considers the use of letter visual information at a single point in time. This is similar to most word recognition models (Norris, 2013). A few models describe accumulation of information over time (Adelman, 2011; Norris, 2006; Norris & Kinoshita, 2012; Norris et al., 2010) and are thus more realistic. In our model we assume that at the end of the 250 ms word presentation the system will have access to the totality of the available position and identity letter information and that a lexical access will be triggered based on this information. We then exclude the possibility of successive lexical top-down feedbacks that might theoretically (despite controversies: D. Norris, McQueen, & Cutler, 2000; Twomey, Kawabata Duncan, Price, & Devlin, 2011) improve letter and word recognition performance.[5]

General Principles: A three-step identification process

Experiment 1 allowed us to quantify visual information extracted in one 250 ms fixation by each observer at a letter level. This information concerns the identity and the position uncertainties of crowded letters located at different eccentricities along the horizontal median. As indicated in **Figure 6**, we hypothesized that the identification and the localization of letters are two independent processes. Following the results of Pelli et al (2003), our model makes the strong assumption that the letters of a word are separately identified before the word recognition step. In consequence, the first step of word recognition is a letter identification first-guess for all letters of the word (what Pelli et al (2003) called a "tentative internal letter identification", a plausible mechanism to explain their experimental results).

Subsequently, the lexical competition occurs at a letter level and is an optimal word discrimination based on this first-guess letter identification step. Many studies suggested that a lexical access based on identity and position letter uncertainties is a plausible process (Davis, 2010; Norris & Kinoshita, 2012; Norris et al., 2010). Here the uncertainty concerning letter identity directly depends on the first

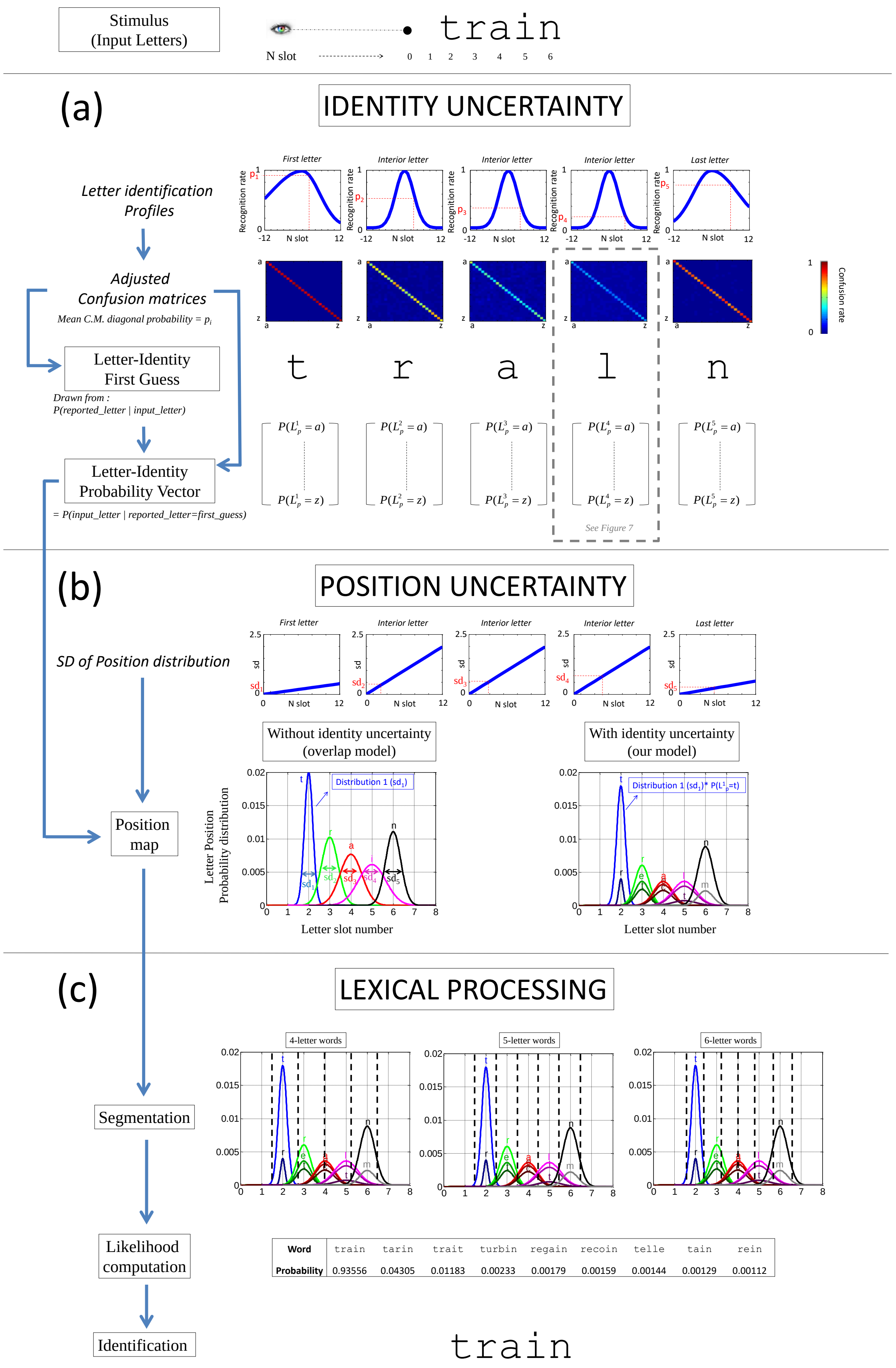
[5] Note that we do not suggest here that no successive lexical feedback occurs during the word recognition process. We only suggest that such loops might have light effects on word recognition performance for a 250 ms fixation task.

identification guess. Our second stage consists in quantifying these identity and positional uncertainties for each letter. To do so, we consider that the model has an internal representation of its confusion errors for any letter first-guess identity, at any letter position and for any relative configuration (i.e. a confusion matrix for each position and configuration case). This knowledge is used to compute a 26-probability vector that represents the probability of any of the 26 letters of the alphabet to be displayed at the letter slot by using the prior letter recognition guess (G. E. Legge et al., 2001; Pelli et al., 2003). This vector is called the letter-identity probability vector (see **Figure 6a** and **Figure 7**). Position uncertainty represents the retinal and neuronal approximation for the spatial localization of letters (Chung & Legge, 2009; Gomez et al., 2008) and is represented by a Gaussian distribution that changes in function of visual eccentricity (see **Figure 6b** for spatial distribution examples). Following the concepts of the overlap model (Gomez et al., 2008; Norris et al., 2010) and the previous theories from whom it was inspired (Krueger, 1978; Ratcliff, 1981), we consider that this spatial distribution describes the perceived spatial locations of letter features before they are grouped in a single object (a letter). In consequence it is important to note the possibility for a single letter to influence the identity of different letter slots.

As shown in **Figure 6c**, the third stage of our model assumes that human readers make optimal use of both identity and position letter-level uncertainties when they try to identify written words. In other words we assume that human readers calculate exact word probabilities from letter probabilities during lexical access. This optimal use of letter probabilities following Bayes' theorem has been suggested by different studies in visual word recognition (Legge et al., 2001; Norris, 2006; Pelli et al., 2003).

Detailed mechanisms

In **Figure 6** we illustrated with an example the principles of our ideal-observer model, i.e. how letter identity and position uncertainty are taken independently into account in order to identify a single word, here the 5-letter word 'train'. As our goal is to predict word recognition performance for each subject individually, we used the individual letter recognition profiles obtained from Experiment 1a and 1b to quantify identity and position uncertainties for each different subject. Note that letter recognition profiles are very different if the observed letter is an exterior letter (data from Experiment 1a) or an interior letter (data from Experiment 1b).



Stimulus
(Input Letters)
train
N slot
0 1 2 3 4 5 6
(a)
IDENTITY UNCERTAINTY
Letter identification Profiles
First letter
Interior letter
Interior letter
Interior letter
Last letter
Recognition rate
N slot
Adjusted Confusion matrices
Mean C.M. diagonal probability = p_i
Confusion rate
Letter-Identity First Guess
Drawn from : P(reported_letter | input_letter)
t r a l n
Letter-Identity Probability Vector
= P(input_letter | reported_letter=first_guess)
See Figure 7
(b)
POSITION UNCERTAINTY
SD of Position distribution
Without identity uncertainty (overlap model)
With identity uncertainty (our model)
Distribution 1 (sd_1)
Distribution 1 (sd_1)* $P(L^1_p=t)$
Letter Position Probability distribution
Letter slot number
Position map
(c)
LEXICAL PROCESSING
4-letter words
5-letter words
6-letter words
Segmentation
Likelihood computation
Identification
Word | train | tarin | trait | turbin | regain | recoin | telle | tain | rein
Probability | 0.93556 | 0.04305 | 0.01183 | 0.00233 | 0.00179 | 0.00159 | 0.00144 | 0.00129 | 0.00112
train

Figure 6: Model principles

The figure shows the different steps of our word recognition ideal observer trying to identify the 5-letter word 'train' displayed parafoveally (centered at slot +4):

(a) Identity uncertainties for the 5 displayed letters are defined as 26-element identity probability vectors (last row). These vectors are obtained based on identity first guesses for each of the five letters (first guess for the different observed letters : 'traln'). These first guesses have been made using letter identification profiles (row 1) and confusion matrices (row 2) that correspond to the recognition rate for each letter position. For instance, the first profile in row 1 corresponds to the "x" blue profile in Figure 3: Exp 1a.

(b) Position uncertainties for the 5 different letters are displayed on position maps (last row). On the left, the position map represents the classical version of the overlap model when identity uncertainty is not coded. Standard deviations on this map are obtained from the linear relationship between eccentricity and standard deviation (row 1). On the right, identity and position uncertainties are coded for each letter and represent the position map for our model. Standard deviations on this map are obtained from the linear relationship between eccentricity and standard deviation, and from the identity uncertainties obtained in (a).

(c) The lexical access is done by using the *position map* to compute the probability of each word of the lexicon to be presented. The position map can be segmented in different number of letters (4, 5, or 6 letters), to calculate word probabilities for each word of the lexicon. Finally, the word with the highest probability is the choice of our ideal observer.

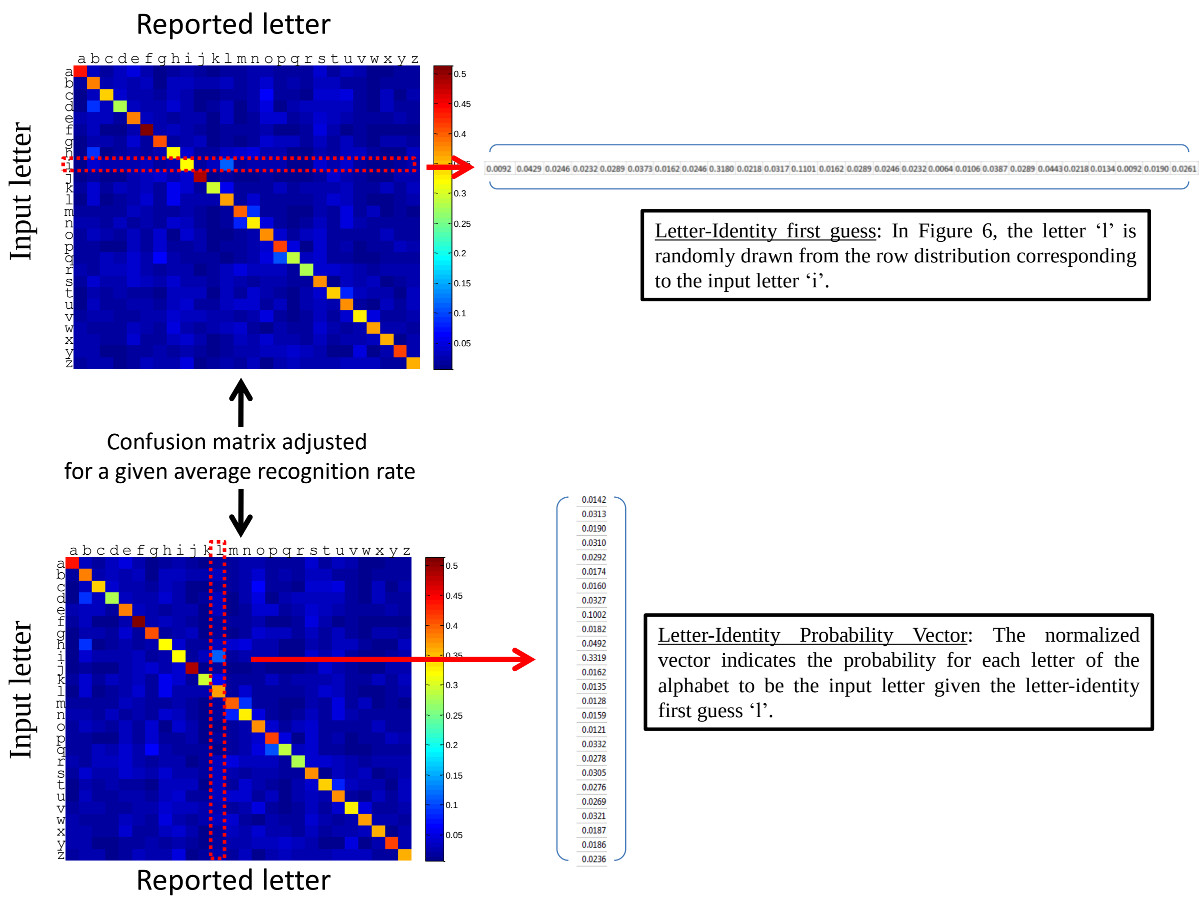


Figure 7 : Confusion matrices and Identity uncertainty

The figure shows how to obtain the *letter-Identity Probability Vector for a letter displayed at a given slot (cf.* Figure 6*a)*. First, a confusion matrix (CM) is calculated so that the average of the values for the diagonal confusion matrix is identical to the letter recognition rate determined by the *letter recognition profile* (see Figure 6a). Here, the corresponding recognition rate is p = 0.37 and the corresponding adjusted confusion matrix is calculated (see Appendix 4). The *first-guess letter* (Pelli et al, 2003) is obtained by randomly drawing from the row distribution corresponding to the displayed letter (Here, the first-guess is a 'l', whereas the letter 'i' was displayed). Finally, the *letter-Identity Probability Vector* is the normalized column corresponding to the first-guess letter, here the letter 'l'.

a. Identity uncertainty coding

Identity uncertainty is coded by a 26-element probability vector for each letter slot. Each element of the vector represents the probability for each letter of the alphabet (a-z) to be the letter at the given slot. Following Pelli et al (2013) evidence, this probability vector is directly based on a first guess of the letter identity.

Letter identity First-Guess

The way for our model to obtain first-guess letters for the different letters of a word is described in **Figure 6a** and with more details on the top of **Figure 7**. To determine the first-guess for each letter, our model uses letter recognition rates and confusion matrices that correspond to each letter slot. Letter recognition profiles that have been measured in Experiments 1a and 1b give the recognition rates for the different letters of the word based on their visual hemifield, their visual eccentricity and their letter relative configuration (left-side crowded only, right-side crowded only, or both-side crowded). The confusion matrix for each letter slot is directly obtained from its letter recognition rate. It is done by transforming a general confusion matrix (obtained with the totality of trials of the five subjects, average recognition rate: 0.44) so that the average recognition rate of the new confusion matrix (the average of its diagonal values) becomes identical to the recognition rate for the specific slot (see the transformation method in Appendix 4). For instance, if the letter identification recognition for a given letter slot position is 0.2 then a confusion matrix that corresponds to 0.2 successful recognition rate and 0.8 confusion rate is firstly built. Secondly, this confusion matrix is used by the ideal observer who uses the letter identity distribution given by the row of the input letter to determine the letter first-guess[6].

Letter Identity Probability Vector

[6] Presented letters correspond to the rows in our confusion matrix examples, and reported letters correspond to the columns.

Once we determine the letter-identity first guess $L_{id}^{s^n}$ for each letter observed at the slot $s^n (n \in [1{:}\, n_{max}]$ for a word with n_{max} letters), we can use Bayes' theorem to define the letter identity uncertainty given this guess, i.e. the probability $P_{id}\left(L_i^{s^n} \middle| L_{id}^{s^n}\right)$ for a letter $L_i^{s^n}$ $(i \in [1{:}\, 26])$ to be the input letter given the first letter identification guess $L_{id}^{s^n}$:

(1) $$P_{id}\left(L_i^{s^n} \middle| L_{id}^{s^n}\right) = \frac{P_{id}\left(L_i^{s^n}\right) * P_{id}\left(L_{id}^{s^n} \middle| L_i^{s^n}\right)}{\sum_j P_{id}\left(L_i^{s^n}\right) * P_{id}\left(L_{id}^{s^n} \middle| L_j^{s^n}\right)}$$

The initial probability $P_{id}\left(L_i^{s^n}\right)$ is the probability of a letter $L_i^{s^n}$ to be displayed at the slot s^n without any visual information and is computed based on the frequency and the orthographic composition of each word of the lexicon with a given word length n_{max}. The likelihood $P_{id}\left(L_{id}^{s^n} \middle| L_i^{s^n}\right)$ is the probability to identify the letter $L_{id}^{s^n}$ given the letter $L_i^{s^n}$ was presented. This probability can be directly obtained by our model because we assume that our ideal observer is aware of the errors it makes for a given slot position after a first letter guess (i.e. the observer knows its own confusion matrices). In this case letter-identity probability vector for the 26 letters of the alphabet corresponds to the column of the confusion matrix for the given letter guess as described in the second part of **Figure 7**.

b. Position uncertainty coding

Figure 6b describes the way we code letter position uncertainty in our model for each letter of the word. It is directly inspired by the overlap model (Gomez et al., 2008) and other studies (Chung & Legge, 2009; Michel & Geisler, 2011) that represents the perceived positions of letters or objects as normal distributions centered on their real positions. In our model, these distributions depend on two parameters: the slot position, and the degree of letter crowding (based on the location and the number of flankers). We used the mislocation coefficients from Experiment 1b (Recognition of letter pentagrams), because the pentagram stimuli were more similar in terms of number of letters to the words presented in Experiment 2. For each condition, Gaussian standard deviations are obtained by using the corresponding mislocation coefficients calculated in Experiments 1a and 1b. In the overlap model, and as shown in the right part of **Figure 6b**, letter identity uncertainty is not taken into account and only letter position is coded. Contrary to our study, where we already measured letter position uncertainty for our subjects in Experiment 1 in order to use it in Experiment 2, the overlap model (Gomez et al., 2008) put the standard deviation value as a free parameter for each letter slot. In the classical version of the overlap model, it produces what we call a *position map* (left panel in **Figure 6b**). It is the overlap between this position map and the position map of any word of the lexicon that defines a similarity measure between words. Mathematically, in the overlap model case, the position map describes the probability P_{pos} of each letter $L_i^{s^n}$ (from a letter slot s^n, $i \in [1{:}26]$) to be presented at a spatial position j as:

(2)
$$\begin{cases} P_{pos}\left(L_i^{s^n}, j\right) = \mathrm{N}^{s^n, \delta^{s^n}}(j) \text{ if } \mathrm{L_i} \text{is the presented letter at slot n} \\ P_{pos}\left(L_i^{s^n}, j\right) = 0 \qquad \text{if } \mathrm{L_i} \text{is not the presented letter at slot n} \end{cases}$$

where $\mathrm{N}^{s^n, \delta^{s^n}}$ is a normal distribution centered on the slot s^n with a corresponding standard deviation δ^{s^n}. Finally, the total probability for each letter L_i ($i \in [1{:}26]$) to be presented at an interval slot $J = [j_1\, j_2]$ would be defined as:

(3)
$$P_{pos}(L_i, J) = \frac{\int_{j_1}^{j_2} \sum_{\mathrm{n} \in [1:\mathrm{n_{max}}]} P_{pos}(L_i^{s^n}, j)\, \mathrm{dj}}{\sum_{\mathrm{i} \in [1:26]} \int_{j_1}^{j_2} \sum_{\mathrm{n} \in [1:\mathrm{n_{max}}]} P_{pos}(L_i^{s^n}, j)\, \mathrm{dj}}.$$

The difference of our model with the overlap model is that we added identity uncertainty to position uncertainty before calculating similarity between words. Position uncertainty was applied to the 26-element identity probability vectors corresponding to each letter slot. In consequence, each letter of the alphabet L_i ($i \in [1{:}26]$) has a probability of presence $P_{id}\left(L_i^{s^n} \middle| L_{id}^{s^n}\right)$ that we name for simplification $P_{id}\left(L_i^{s^n}\right)$ for each letter slot s^n. In the right part of the **Figure 6b**, we modified the overlap model version by combining in an independent way identity and position uncertainty values to obtain the probability $P_{pos}(L_i^{s^n}, j)$ of each letter $L_i^{s^n}$ (from a letter slot s^n) to be presented at a spatial position j : $P_{pos}(L_i^{s^n}, j) = P_{id}\left(L_i^{s^n}\right) * \mathrm{N}^{s^n, \delta s^n}(j)$. Finally, the total probability for a letter L_i to be presented at an interval slot $J = [j_1\, j_2]$ is similarly defined as:

(4) $$P_{pos}(L_i, J) = \frac{\int_{j_1}^{j_2} \sum_{\mathrm{n} \in [1:\mathrm{n_{max}}]} P_{pos}(L_i^{s^n}, j)\, \mathrm{dj}}{\sum_{\mathrm{i} \in [1:26]} \int_{j_1}^{j_2} \sum_{\mathrm{n} \in [1:\mathrm{n_{max}}]} P_{pos}(L_i^{s^n}, j)\, \mathrm{dj}} = \frac{\int_{j_1}^{j_2} \sum_{\mathrm{n} \in [1:\mathrm{n_{max}}]} P_{id}\left(L_i^{s^n}\right) * \mathrm{N}^{s^n, \delta s^n}(j)\, \mathrm{dj}}{\sum_{\mathrm{i} \in [1:26]} \int_{j_1}^{j_2} \sum_{\mathrm{n} \in [1:\mathrm{n_{max}}]} P_{id}\left(L_i^{s^n}\right) * \mathrm{N}^{s^n, \delta s^n}(j)\, \mathrm{dj}}$$

c. Lexical access coding

The lexical access is described in Figure 7c. It follows the same multiplication principle as in (Legge et al., 2001) and in (Norris, 2006). Posterior probability is computed for each word w of the lexicon following the equation:

(5) $$P(w|(L_{id}^{s^1}, \ldots, L_{id}^{s^{max}})) = \frac{P(w) * P((L_{id}^{s^1}, \ldots, L_{id}^{s^{max}})\big|w)}{\sum_{w=1:w_{max}} P((L_{id}^{s^1}, \ldots, L_{id}^{s^{max}})\big|w}$$

The difference with the lexical access described in (Legge et al., 2001) is that in order to identify words of different length compared to the presented word, our lexical access also follows the principles of the overlap model: The likelihood value $P((L_{id}^{s^1}, \ldots, L_{id}^{s^{max}})\big|w)$ for a word w of the lexicon is calculated by segmenting the position map in spatial equal parts $J_1 \ldots J_{l_{max}}$ based on the number of letters of $w = \{l_1 \ldots l_{l_{max}}\}$:

(6) $$P\left((L_{id}^{s^1}, \ldots, L_{id}^{s^{max}})\Big|w\right) = \prod_{i=1:lmax} P((L_{id}^{s^1}, \ldots, L_{id}^{s^{max}}) \subset J_i \Big| l_i)$$

We can write that:

$$(7)\qquad P((L_{id}^{s^1},\dots,L_{id}^{s^{max}})\subset J_i \middle| l_k) = P_{pos}(l_k, J_i = [j_i\, j_{i+1}]) = \frac{\int_{j_i}^{j_{i+1}} \sum_{\mathrm{n}\in[1:\mathrm{n_{max}}]} P_{id}\left(l_k^{s^n}\right) * \mathrm{N}^{s^n,\delta s^n}(j)\,\mathrm{dj}}{\sum_{\mathrm{i}\in[1:26]} \int_{j_i}^{j_{i+1}} \sum_{\mathrm{n}\in[1:\mathrm{n_{max}}]} P_{id}\left(l_k^{s^n}\right) * \mathrm{N}^{s^n,\delta s^n}(j)\,\mathrm{dj}}$$

and finally:

$$(8)\qquad P\left((L_{id}^{s^1},\dots,L_{id}^{s^{max}})\middle| w\right) = \prod_{k=1:lmax} \frac{\int_{j_i}^{j_{i+1}} \sum_{\mathrm{n}\in[1:\mathrm{n_{max}}]} P_{id}\left(l_k^{s^n}\right) * \mathrm{N}^{s^n,\delta s^n}(j)\,\mathrm{dj}}{\sum_{\mathrm{i}\in[1:26]} \int_{j_i}^{j_{i+1}} \sum_{\mathrm{n}\in[1:\mathrm{n_{max}}]} P_{id}\left(l_k^{s^n}\right) * \mathrm{N}^{s^n,\delta s^n}(j)\,\mathrm{dj}}$$

The prior probability $P(w)$ for a word w from the lexicon is defined as the product of the logarithm of the word frequency and another probability based on both the length of the word of the lexicon and the presented word:

$$(9)\qquad P_{initial}(w) = \log\left(frequency(w)\right) * P(l_{max} = n_{max})$$

The logarithm of the word frequency is typically used in word recognition models (Engbert, Nuthmann, Richter, & Kliegl, 2005; Reichle, Rayner, & Pollatsek, 2003) to take into account the effect of word frequency on word recognition performance for a given duration (Howes & Solomon, 1951; Whaley, 1978). The frequency probabilities used in our model come from the Lexique corpus for French lemma words (New et al., 2004). The length-difference probability represents the probability for a subject to report a word that does not have the same length as the displayed word (whereas before the trial the subject knew the number of letters of each displayed word given our experimental word-length blocked design). Probabilities for a word of the lexicon with a certain length to be reported based on the length of the displayed word are given in Appendix 5 for the different subjects and are directly used by the ideal observer. Because more than 95% of errors are 1-letter length errors for our subjects, we considered that the only possible word length-errors in our model are one-letter length errors. These word length errors usually represent between 5% and 20% of the errors, depending on the subjects and the experimental conditions. Finally our ideal observer decides that the word which is displayed is the word with the highest calculated probability.

COMPARISONS WITH WORD RECOGNITION DATA FOR EACH SUBJECT

For our simulations we asked our model to identify exactly the same word stimuli that had been identified by each observer (same word identity and same location). Each word was identified 512 times by our model, and we defined the corresponding average recognition rate as the model performance of the model for this word. For each simulated observer we created a particular ideal observer that was constrained by the same letter position and identity limitations as the real observer: For each subject letter identity uncertainty was defined based on the corresponding letter recognition profiles. In addition letter position uncertainty was defined based on the corresponding mislocation coefficients (Experiment 1a for single-side crowded letters and Experiment 1b for double-side crowded letters). The model was implemented in Matlab (Mathworks, Inc), and we saved the word recognition errors made by the five different ideal observers during the simulations. For each observer we compared human and model performance to assess if they matched for the effect of eccentricity on word recognition performance and for corresponding letter recognition errors.

Word recognition errors

Word recognition rates for behavioral and model data are shown in **Figure 5a** for each eccentricity, each word length, each subject, and on average across subjects. The corresponding scatterplots are shown in **Figure 8a**. Each point corresponds to a condition with a given eccentricity and a given word length for all subjects. They show a pretty good correlation between both values. A mixed-effects analysis was run to test the relationship between the experimental word recognition performance and the word recognition performance predicted by our model. The dependent variable was the model values; the fixed variables were the eccentricity, the word length and the experimental values; the random variable was the subject variable. Results of the mixed-effects analysis are shown in Table 3a: We indicate the estimated values of the intercept and of variable coefficients. We also indicate 95% confidence intervals for these values. Our goal was to estimate how far these values are from a 0 value for the intercept and from a 1 value for the experimental data coefficient as would be the case if our model had perfect predictions. Statistical results show an intercept value of 0.04 (CI: [-0.06; 0.14] and a coefficient value of 0.81 for the experimental data (CI: [0.75; 0.88]), thus suggesting an accurate prediction.

Fixed effects:

(a)	Estimate	Std. Error	t value	CI 2.5%	CI 97.5%
(Intercept)	0.04	0.05	0.72	-0.06	0.14
Experimental rates	0.81	0.03	24.62	0.75	0.88
Eccentricity	- 0.03	0.02	-2.10	- 0.07	0.00
Word length	0.01	0.01	1.79	0.00	0.02

(b)	Estimate	Std. Error	t value	CI 2.5%	CI 97.5%
(Intercept)	0.07	0.04	1.57	-0.03	0.16
Experimental rates	0.88	0.03	25.79	0.81	0.94
Eccentricity	-0.03	0.01	-2.46	-0.05	0.00
Word length	0.00	0.00	0.55	-0.01	0.01

(c)	Estimate	Std. Error	t value	CI 2.5%	CI 97.5%
(Intercept)	-0.12	0.05	-2.43	-0.21	-0.02
Experimental rates	1.06	0.05	21.72	0.96	1.15
Eccentricity	-0.02	0.01	-2.23	-0.04	0.00
Word length	0.00	0.00	2.07	0.00	0.01

Tableau 3 : Fixed effects results of the linear mixed-effects models to predict model recognition rates in function of experimental letter recognition rates, eccentricity and word length as shown in Fig. 8. (a) Word recognition performance, (b) Letter within-word performance (considering mislocation and identity errors), (c) Letter within-word errors (considering identity errors only). The last two columns show the lower and upper values of 95% confidence intervals. Note the different Y-axis scales in (b) and (c)

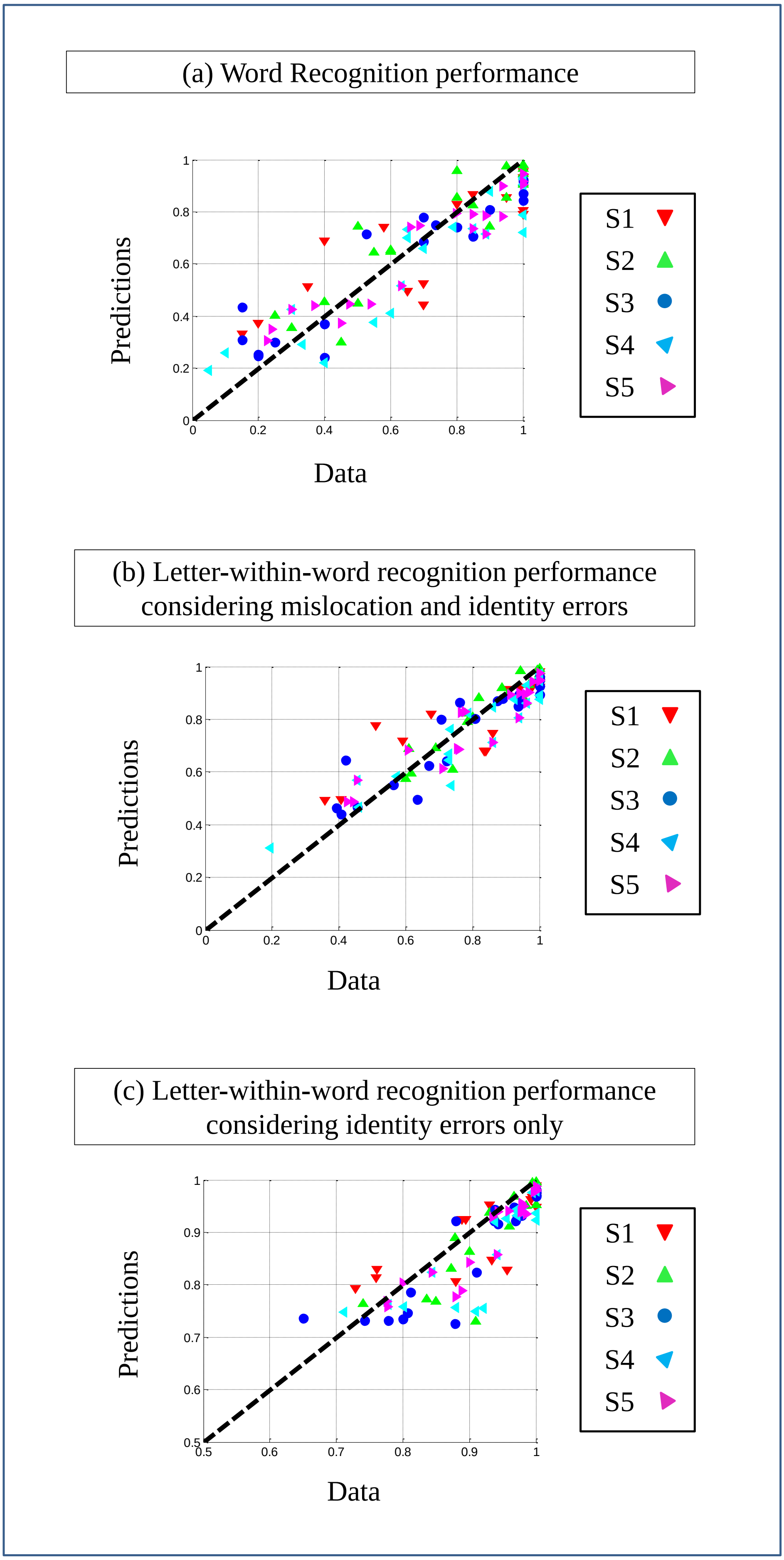
We also predicted average word recognition performance based on average visual span profiles and average mislocation coefficients (see **Figure 5a**). The corresponding correlation coefficient is very high (r^2 = 0.92).

Letter recognition errors

To compare the letter errors made by the subjects with the letter errors made by the corresponding models during the word recognition task, we considered two kinds of errors: (case 1) an error was defined either as an identity error or as a mislocation error and (case 2) an error was defined as an identify error only (i.e. a correctly identified letter was not considered as an error if its location was not correctly reported). For instance, if the word 'balle' was reported when the word "table" was displayed, 3 letters were correctly identified and reported at their correct position ("_a_le") which was counted as two errors in case 1 above. However, the reported letter "b" was not considered as an error in case 2 above as this letter was displayed in "table", so that only one error occurred in this case.. These analyses characterize the letter errors that are made by observers during word recognition at a confusion level and at a mislocation level. For the two cases described above, **Figure 5b** and **Figure 5c** show the experiment and model data for each eccentricity, each word length, each subject, and on average across subjects. The corresponding scatterplots are shown in **Figure 8b** and **Figure 8c**. Each point corresponds to a given eccentricity and word length condition. Two Mixed-effects analyses were run to test the relationship between experimental and simulated letter within-word recognition performance. The dependent variable was the model values for letter recognition; the fixed variables were the eccentricity, the word length and the experimental data; the random variable was the subject variable. Results of the mixed-effects analyses are shown in Table 3b and 3c. For letter within-word performance considering mislocation and identity errors (Fig. 8.b), results suggest an accurate fit between experiment and model data. For letter within-word performance considering identity errors only (Fig. 8.c), results suggest that the model slightly underestimates experimental results.

We also predicted average letter recognition performance based on average visual span profiles and average mislocation coefficients (see **Figure 5b and 5c**). The corresponding correlation coefficient were also very high (r^2 = 0.92 and 0.86 for both types of errors).

Overall, these results suggest that ideal observers make similar errors to those made by real subjects at a word- and letter-level when they try to identify words presented at foveal and parafoveal locations. However, part (c) of Figure 8 and our statistical analysis show that the model tends to over-estimate the number of identity errors as letter recognition is under-estimated. This observation is commented in the Discussion section.



(a) Word Recognition performance
Predictions
Data
S1
S2
S3
S4
S5
(b) Letter-within-word recognition performance considering mislocation and identity errors
Predictions
Data
S1
S2
S3
S4
S5
(c) Letter-within-word recognition performance considering identity errors only
Predictions
Data
S1
S2
S3
S4
S5

Figure 8 : Scatterplots for word recognition rates comparing experimental data and model predictions

The scatterplots show the comparison between experimental data and model predictions for each presentation condition (5 word eccentricities x 3 word lengths), for each of the 5 subjects

DISCUSSION

The novelty of our study is the use of an accurate and individual behavioral quantification of letter identity and letter position uncertainties to test and validate a three-stage model of foveal and parafoveal word recognition.

Letter identity uncertainty is usually skipped or represented in a very simplistic way in word recognition models (Norris, 2013). Here we carefully took into account the joint effects of visual acuity and visual crowding on word letters by measuring letter identification profiles and confusion matrices for each subject in experiments excluding word context. We also used these data to represent letter position uncertainty by normal distributions centered on each actual letter position (Chung & Legge, 2009) for each subject. Our model is thus unique in its accurate specification of the two types of perceptual uncertainties that limit letter recognition performance across the visual field: identity and position uncertainties (Davis, 2010; Norris & Kinoshita, 2012; Yu et al., 2014), known to be exacerbated in crowded conditions (Harrison & Bex, 2016; van den Berg et al., 2012).

Knowing these uncertainties with a great accuracy allowed us to implement and test a foveal and parafoveal parameter-free word recognition model based on the following 3 stages: (1) An independent and parallel letter processing stage that precedes word identification as shown in (Pelli et al., 2003) and provides “letter guess” stage, (2) a stage where identity and positional uncertainties are calculated based on letter first guesses and (3) a lexical access stage where both letter identity and position uncertainties are combined together in order to identify written words. Comparing word recognition performance between five human and five simulated observers with corresponding letter position and identity uncertainties tend to corroborate our assumption of this 3-stage word-recognition model (**Figure 8a, 8b, and 8c).**

These results confirm previous assumptions concerning the links between foveal/parafoveal letter recognition and word recognition performance (Legge et al., 2007; Nazir et al., 1992; Stevens & Grainger, 2003). One of the strongest assumptions about this link was made by Pelli et al (Pelli et al.,

2003) who suggested that humans identify foveal words (in a low contrast and noisy environment) in a non-optimal way, automatically identifying letter components of words (i.e. observers would attempt to “guess” letters before “guessing” words) before to consider the word recognition stage. The performance of our simulated observers compared to expert human readers is a strong support to this theory when letter visibility is degraded by visual eccentricity (rather than by visual noise). In supplement, we also suggest that if expert readers are not optimal in the way they identify words because of this first automatic letter guess, they look optimal in the way they use this information to identify foveal or parafoveal words: First, expert readers would improve their first letter guesses by transforming them in letter identity uncertainties using their knowledge of the confusion errors that they make when they try to identify crowded letters in the periphery (i.e. a knowledge of different confusion matrices based on the letter locations). Second, expert readers would also perfectly integrate the obtained letter identity uncertainties and the corresponding positional uncertainties. The optimal correction of the first guess is a plausible hypothesis given the billions of crowded letters recognized during expert readers' life in order to identify foveal and parafoveal words. The natural reading process, where a parafoveal word is first previewed and then usually foveated, is a perfect task for the visual system to learn the perceptual errors made while automatically identifying parafoveal letters. Finally, it is remarkable that adding another optimal stage for the use of both uncertainties during the lexical access makes the performance and the errors of the model so similar to human reader performance.

Our study also suggests two important facts: (1) The amount of letter information extracted by a human reader during a 250 ms fixation is the same regardless of the relative position of a letter within a word (note that the within-word relative position must not be confused with the relative position within the interior trigram as used in experiments 1a and 1b). This suggests that there is no special advantage for letters located at a certain relative position within a word, only absolute letter eccentricity, hemifield and number of flanking letters seem to influence letter identity and position uncertainties (2) the amount of information extracted for each letter within a word seems independent of the word length: The amount of extracted information for each individual letter (for a given configuration and eccentricity) could be similar if the word is a 5-letter, 7-letter, or 9-letter word or if 5 random letters are simultaneous reported among a letter string. These findings are a strong support to theories advocating that letter processing is parallel and independent of letter-cluster or word contexts during the foveal or parafoveal identification of words (Adelman, Marquis, & Sabatos-DeVito, 2010; Davis, 2010). In addition to being mutually independent and parallel during word recognition, multi-letter recognition processing does not seem to be constrained by a form of short term visual memory (Castet, Descamps, Denis-Noël, & Colé, 2017), a direct support to the theory that letter recognition could not be similar to object recognition for expert readers.

While word recognition performance and letter errors seem overall correctly predicted by our model, some differences exist when we pay attention to the patterns of letter errors made by the subjects (Figure 8c). Indeed, our model under-estimates letter-within-word recognition performance by overestimating the number of identity errors. One possible cause explaining this result may be the way we simulated the confusion matrices (called CM_{new} in the appendix) in our model in order to quantify identity uncertainty for each eccentricity. The computation of these new confusion matrices is based on a single confusion matrix (called CM_{all} in the appendix) obtained by combining the letter identification errors from all 5 participants in the letter recognition tasks. Thus, any new confusion matrix does not take into account individual differences. Moreover, in a subsequent step, any new confusion matrix results from an interpolation between the CM_{all} matrix (average recognition rate = 44%) and one of two extreme matrices (either a 100% accuracy matrix or a chance-level matrix – see details in appendix). The estimated letter errors made by our model could in consequence be less accurate than human ones, leading to more letter identity errors. This could explain why the letter identity errors of the model are larger for far eccentricities (see Figure 4, where most of the model errors occur beyond +/- 6 letter slots). To solve this problem, a better method would be to extract behavioral confusion matrices based on letter recognition tasks of different difficulties/eccentricities as done in Legge et al, 2001. However, such measurements necessitate an extremely large set of data that we did not have in this study. Another non-exclusive explanation of this difference could be due to the word superiority effect (Reicher, 1969) and the fact that letters could be better recognized in words compared to random strings. On another hand, letter mislocations could have been under-estimated by our measurements based on strings of three or five letters.

What are the alternative models that could be compared to the one presented here? Most of the implemented word recognition models skip the first letter guess demonstrated by Pelli et al (2003), considering the use of available letter identity information as optimal during the lexical decision step. This is the case for the interactive activation model (McClelland & Rumelhart, 1981) and its numerous inspired models. Here, we have the possibility to simulate a model that would skip Pelli's letter first guess and would consider that the available letter identity information is extracted then used in an optimal way by human readers. The results of this "optimal" model are shown in Figure 9 and compared to our experimental and model data. It is very interesting to see that this model would predict a recognition rate of almost 100% for word lengths of 5, 7 and 9 letters in all conditions of our experiments, which is clearly overestimating our experimental data. This result tends to confirm the results of Pelli et al (2003) that word recognition cannot be considered as a process making an optimal use of the available visual information: It suggests that a letter identification step automatically precedes word recognition and should theoretically be a part of any word recognition model.

In Figure 9, we also show another version of our model that would consider only letter identity uncertainty and would skip letter position uncertainty. It is the equivalent of the Legge et al model (Legge et al., 2001) applied to single word recognition. Results for the three word lengths suggest the importance of letter position uncertainty in peripheral word recognition errors: Removing letter position uncertainty from our original model decreases the error rates by at least 50% for each condition. This result suggests that letter position uncertainty is an important and overlooked factor limiting peripheral word recognition and reading without central vision in general.

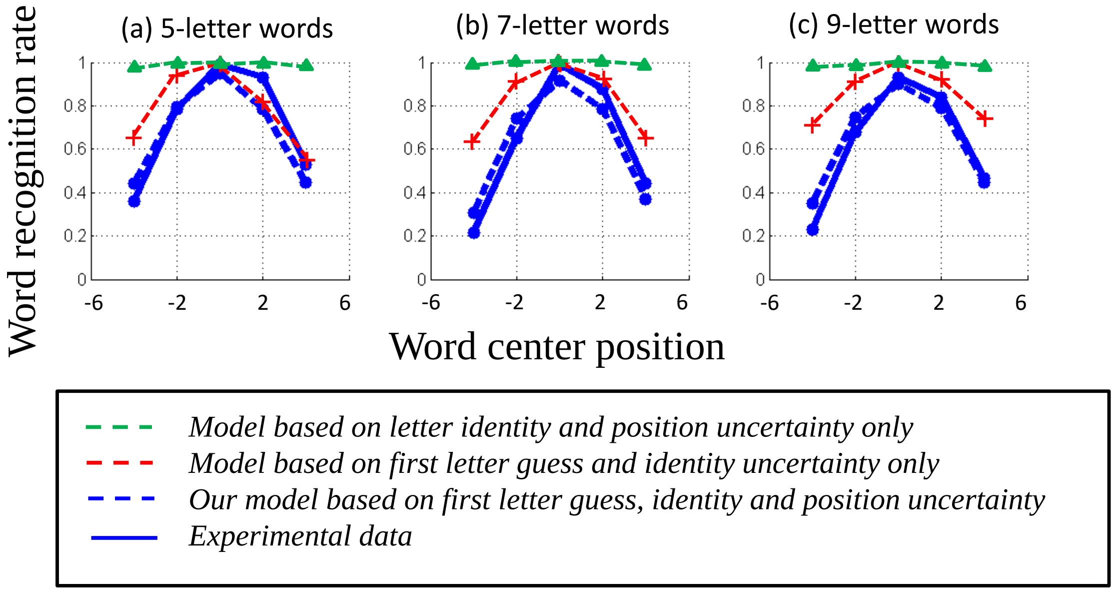


Figure 9 : Comparison of different models with our experimental data. Average word recognition rate is plotted as a function of word center position for behavioral (solid lines) and model data (dashed lines). The three plots correspond to the three word lengths used in our experiments (a) 5-letter, (b) 7-letter and (c) 9-letters.

Finally, there are multiple potential applications of our work. First, our results suggest that a part of visual limitations in word recognition (and probably in natural reading as well) can be characterized by simple measurements of letter identity and position uncertainties. This corroborates recent studies that link letter recognition and reading performance in normally-sighted readers (Frömer et al., 2015; Risse, 2014), and suggest that part of individual differences (including some forms of dyslexia) in reading can be due to differences in low-level visual factors. Here, we show that simple measurements with letter strings may be sufficient to characterize these low-level factors and predict differences in word recognition performance. How could this model be applied to normal natural reading (i.e. when eyes move in order to identify the words of a sentence) ? The prior word probability in our Bayes equation for lexical processing is based on frequency and word length. It can reflect the level of activation for each word of the lexicon. A simple way to modify our model would be (1) to introduce word predictability in the prior value and (2) to add an equation to calculate the integration between parafoveal and foveal information, for instance following a Bayesian inference as in (Norris, 2006). Like word frequency, word predictability would directly modulate the probability for a word in the lexicon to be presented independently of the available visual information. Predictions of eye movements (duration of fixations and the amplitude of saccades) and their possible interaction with lexical processing would be another step to take into consideration for a complete model (Bicknell & Levy, 2010). A model following similar principles to those of our model (Legge et al., 2002) showed the possibility to predict patterns of fixation locations during reading. In a future version, our model could similarly link fixation duration to the amount of letter information extracted in each fixation to identify words. For people with central field loss, other oculomotor factors (fixation instability, lack of saccade accuracy) or perceptual factors (presence of scotomas) would need to be added.

An interesting application for our model also concerns the possibility to predict parafoveal and peripheral word recognition performance at any position across the visual field based on a possible measurement of parafoveal and peripheral letter recognition performance (not just on the horizontal median as in our experiment). This is of critical importance for patients with central field loss who cannot use their central vision and need to learn to use a new extra-foveal area (Cheung & Legge, 2005). Using such a model could help the definition of extra-foveal areas that would be optimal for the parafoveal and peripheral recognition of words, a question that could be crucial in order to optimize reading performance in patients with central field loss by training them to use these new retinal areas.

Acknowledgements

For this work, Dr Jean-Baptiste Bernard was supported by a BLRI (Brain and Language Research Institute) postdoctoral fellowship and by grants from the Fondation de l'Avenir and Fondation Visaudio. We thank Dr Francoise Vitu-Thibault for her helpful comments, and Dr Fermin Moscoso del Prado Martin for the conception of an initial version of the model.

References

Adelman, J. S. (2011). Letters in time and retinotopic space. *Psychological Review*, *118*(4), 570-582. https://doi.org/10.1037/a0024811

Adelman, J. S., Marquis, S. J., & Sabatos-DeVito, M. G. (2010). Letters in words are read simultaneously, not in left-to-right sequence. *Psychological Science*, *Vol.21*(No.12), 1799-1801.

Balota, D. A., Yap, M. J., & Cortese, M. J. (2006). Visual word recognition: The journey from features to meaning (a travel update). *Handbook of psycholinguistics*, *2*, 285–375.

Bernard, J.-B., & Chung, S. T. (2011). The dependence of crowding on flanker complexity and target–flanker similarity. *Journal of Vision*, *11*(8). Consulté à l’adresse http://www.journalofvision.orgwww.journalofvision.org/content/11/8/1.short

Bicknell, K., & Levy, R. (2010). Rational eye movements in reading combining uncertainty about previous words with contextual probability. *Proceedings of the Annual Meeting of the Cognitive Science Society*, *32*(32). Consulté à l’adresse https://escholarship.org/uc/item/1mm294p6

Bouma, H. (1970). Interaction Effects in Parafoveal Letter Recognition. *Nature*, *226*(5241), 177-178. https://doi.org/10.1038/226177a0

Bouma, H. (1973). Visual interference in the parafoveal recognition of initial and final letters of words. *Vision Research*, *13*(4), 767-782. https://doi.org/10.1016/0042-6989(73)90041-2

Bouma, H., & Legein, C. P. (1977). Foveal and parafoveal recognition of letters and words by dyslexics and by average readers. *Neuropsychologia*, *15*(1), 69-80.

Brysbaert, M., Vitu, F., & Schroyens, W. (1996). The right visual field advantage and the optimal viewing position effect: On the relation between foveal and parafoveal word recognition. *Neuropsychology*, *10*(3), 385-395. https://doi.org/10.1037/0894-4105.10.3.385

Carreiras, M., Armstrong, B. C., Perea, M., & Frost, R. (2014). The what, when, where, and how of visual word recognition. *Trends in Cognitive Sciences*, *18*(2), 90-98. https://doi.org/10.1016/j.tics.2013.11.005

Castet, E., Descamps, M., Denis-Noël, A., & Colé, P. (2017). Letter and symbol identification: No evidence for letter-specific crowding mechanisms. *Journal of Vision*, *17*(11), 2. https://doi.org/10.1167/17.11.2

Cheung, S.-H., & Legge, G. E. (2005). Functional and cortical adaptations to central vision loss. *Visual neuroscience*, *22*(2), 187-201. https://doi.org/10.1017/S0952523805222071

Chung, S. T., & Legge, G. E. (2009). Precision of position signals for letters. *Vision research*, *49*(15), 1948–1960.

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*(1), 204-256. https://doi.org/10.1037/0033-295X.108.1.204

Davis, C. J. (2010). The spatial coding model of visual word identification. *Psychological Review*, *117*(3), 713-758. https://doi.org/10.1037/a0019738

Engbert, R., Nuthmann, A., Richter, E. M., & Kliegl, R. (2005). SWIFT: a dynamical model of saccade generation during reading. *Psychological Review*, *112*(4), 777-813. https://doi.org/10.1037/0033-295X.112.4.777

Fiset, D., Blais, C., Ethier-Majcher, C., Arguin, M., Bub, D., & Gosselin, F. (2008). Features for identification of uppercase and lowercase letters. *Psychological Science*, *19*(11), 1161–1168.

Freeman, J., Chakravarthi, R., & Pelli, D. G. (2012). Substitution and pooling in crowding. *Attention, Perception & Psychophysics*, *74*(2), 379-396. https://doi.org/10.3758/s13414-011-0229-0

Frömer, R., Dimigen, O., Niefind, F., Krause, N., Kliegl, R., & Sommer, W. (2015). Are Individual Differences in Reading Speed Related to Extrafoveal Visual Acuity and Crowding? *PLoS ONE*, *10*(3), e0121986. https://doi.org/10.1371/journal.pone.0121986

Geisler, W. S. (1989). Sequential ideal-observer analysis of visual discriminations. *Psychological Review*, *96*(2), 267-314.

Geisler, Wilson S. (2011). Contributions of Ideal Observer Theory to Vision Research. *Vision research*, *51*(7), 771-781. https://doi.org/10.1016/j.visres.2010.09.027

Gomez, P., Ratcliff, R., & Perea, M. (2008). The overlap model: a model of letter position coding. *Psychological Review*, *115*(3), 577-600. https://doi.org/10.1037/a0012667

Grainger, J., & Jacobs, A. M. (1996). Orthographic processing in visual word recognition: a multiple read-out model. *Psychological Review*, *103*(3), 518-565.

Grainger, Jonathan, Rey, A., & Dufau, S. (2008). Letter perception: from pixels to pandemonium. *Trends in Cognitive Sciences*, *12*(10), 381-387. https://doi.org/10.1016/j.tics.2008.06.006

Hanus, D., & Vul, E. (2013). Quantifying error distributions in crowding. *Journal of Vision*, *13*(4), 17-17. https://doi.org/10.1167/13.4.17

Harrison, W. J., & Bex, P. J. (2016). Visual crowding is a combination of an increase of positional uncertainty, source confusion, and featural averaging. *BioRxiv*, 088898. https://doi.org/10.1101/088898

He, Y., Legge, G. E., & Yu, D. (2013). Sensory and cognitive influences on the training-related improvement of reading speed in peripheral vision. *Journal of Vision*, *13*(7), 14-14. https://doi.org/10.1167/13.7.14

Howes, D. H., & L, R. (1951). Visual duration threshold as a function of word-probability. *Journal of Experimental Psychology*, *41*(6), 401-410. https://doi.org/10.1037/h0056020

Kersten, D., Mamassian, P., & Yuille, A. (2004). Object perception as Bayesian inference. *Annual Review of Psychology*, *55*, 271-304. https://doi.org/10.1146/annurev.psych.55.090902.142005

Krueger, L. E. (1978). A Theory of Perceptual Matching. *Psychological Review*, *85*(4), 278–304.

Latham, K., & Whitaker, D. (1996). A Comparison of Word Recognition and Reading Performance in Foveal and Peripheral Vision *. *Vision Research*, *36*(17), 2665-2674. https://doi.org/10.1016/0042-6989(96)00022-3

Legge, G. E., Mansfield, J. S., & Chung, S. T. (2001). Psychophysics of reading. XX. Linking letter recognition to reading speed in central and peripheral vision. *Vision Research*, *41*(6), 725-743.

Legge, Gordon E., & Bigelow, C. A. (2011). Does print size matter for reading? A review of findings from vision science and typography. *Journal of Vision*, *11*(5). https://doi.org/10.1167/11.5.8

Legge, Gordon E., Cheung, S.-H., Yu, D., Chung, S. T. L., Lee, H.-W., & Owens, D. P. (2007). The case for the visual span as a sensory bottleneck in reading. *Journal of vision*, *7*(2), 9.1-915. https://doi.org/10.1167/7.2.9

Legge, Gordon E., Hooven, T. A., Klitz, T. S., Stephen Mansfield, J. S., & Tjan, B. S. (2002). Mr. Chips 2002: new insights from an ideal-observer model of reading. *Vision Research*, *42*(18), 2219-2234.

Legge, Gordon E., Klitz, T. S., Tjan, B. S., & Low-vision, M. L. F. (1997). Mr. Chips: An ideal-observer model of reading. *Psychological Review*, *104*, 524–553.

Levi, D. M., & Tripathy, S. P. (1996). Localization of a peripheral patch: the role of blur and spatial frequency. *Vision Research*, *36*(23), 3785-3803.

Levi, Dennis M. (2008). Crowding—An essential bottleneck for object recognition: A mini-review. *Vision research*, *48*(5), 635–654.

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), 375-407. https://doi.org/10.1037/0033-295X.88.5.375

Michel, M., & Geisler, W. S. (2011). Intrinsic position uncertainty explains detection and localization performance in peripheral vision. *Journal of Vision*, *11*(1), 18. https://doi.org/10.1167/11.1.18

Najemnik, J., & Geisler, W. S. (2005). Optimal eye movement strategies in visual search. *Nature*, *434*(7031), 387-391. https://doi.org/10.1038/nature03390

Nazir, T. A., Heller, D., & Sussmann, C. (1992). Letter visibility and word recognition: The optimal viewing position in printed words. *Perception & Psychophysics*, *52*(3), 315-328. https://doi.org/10.3758/BF03209148

New, B., Pallier, C., Brysbaert, M., & Ferrand, L. (2004). Lexique 2: A new French lexical database. *Behavior Research Methods, Instruments, & Computers*, *36*(3), 516–524.

Norris, D., McQueen, J. M., & Cutler, A. (2000). Merging information in speech recognition: feedback is never necessary. *The Behavioral and Brain Sciences*, *23*(3), 299-325; discussion 325-370.

Norris, Dennis. (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, Dennis. (2013). Models of visual word recognition. *Trends in Cognitive Sciences*, *17*(10), 517-524. https://doi.org/10.1016/j.tics.2013.08.003

Norris, Dennis, & 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

Norris, Dennis, Kinoshita, S., & van Casteren, M. (2010). A stimulus sampling theory of letter identity and order. *Journal of Memory and Language*, *62*(3), 254-271. https://doi.org/10.1016/j.jml.2009.11.002

O'Reagan, J. K., & Jacobs, A. M. (1992). Optimal viewing position effect in word recognition: A challenge to current theory. *Journal of Experimental Psychology: Human Perception and Performance*, *18*(1), 185-197. https://doi.org/10.1037/0096-1523.18.1.185

Peirce, J. W. (2007). PsychoPy—Psychophysics software in Python. *Journal of Neuroscience Methods*, *162*(1-2), 8. https://doi.org/10.1016/j.jneumeth.2006.11.017

Pelli, D. G., Farell, B., & Moore, D. C. (2003). The remarkable inefficiency of word recognition. *Nature*, *423*(6941), 752-756. https://doi.org/10.1038/nature01516

Pelli, D. G., Palomares, M., & Majaj, N. J. (2004). Crowding is unlike ordinary masking: distinguishing feature integration from detection. *Journal of Vision*, *4*(12), 1136-1169. https://doi.org/10:1167/4.12.12

Pelli, D. G., & Tillman, K. A. (2008). The uncrowded window of object recognition. *Nature neuroscience*, *11*(10), 1129–1135.

Pelli, D. G., Tillman, K. A., Freeman, J., Su, M., Berger, T. D., & Majaj, N. J. (2007). Crowding and eccentricity determine reading rate. *Journal of Vision*, *7*(2), 20.1-36. https://doi.org/10.1167/7.2.20

R Development Core Team. (2013). R Development Core Team (2013). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL http://www.R-project.org. R Foundation for Statistical Computing, Vienna, Austria.

Ratcliff, R. (1981). A theory of order relations in perceptual matching. *Psychological Review*, *88*(6), 552-572. https://doi.org/10.1037/0033-295X.88.6.552

Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin*, *124*(3), 372-422.

Reicher, G. M. (1969). Perceptual recognition as a function of meaningfulness of stimulus material. *Journal of Experimental Psychology*, *81*(2), 275-280. https://doi.org/10.1037/h0027768

Reichle, E. D., Rayner, K., & Pollatsek, A. (2003). The E-Z reader model of eye-movement control in reading: comparisons to other models. *The Behavioral and Brain Sciences*, *26*(4), 445-476; discussion 477-526.

Renninger, L. W., Verghese, P., & Coughlan, J. (2007). Where to look next? Eye movements reduce local uncertainty. *Journal of Vision*, *7*(3), 6. https://doi.org/10.1167/7.3.6

Risse, S. (2014). Effects of visual span on reading speed and parafoveal processing in eye movements during sentence reading. *Journal of Vision*, *14*(8), 11. https://doi.org/10.1167/14.8.11

Stevens, M., & Grainger, J. (2003). Letter visibility and the viewing position effect in visual word recognition. *Perception & Psychophysics*, *65*(1), 133-151.

Strasburger, H. (2005). Unfocussed spatial attention underlies the crowding effect in indirect form vision. *Journal of Vision*, *5*(11), 8. https://doi.org/10.1167/5.11.8

Strasburger, H., & Malania, M. (2013). Source confusion is a major cause of crowding. *Journal of Vision*, *13*(1), 24. https://doi.org/10.1167/13.1.24

Twomey, T., Kawabata Duncan, K. J., Price, C. J., & Devlin, J. T. (2011). Top-down modulation of ventral occipito-temporal responses during visual word recognition. *Neuroimage*, *55*(3-3), 1242-1251. https://doi.org/10.1016/j.neuroimage.2011.01.001

van den Berg, R., Johnson, A., Martinez Anton, A., Schepers, A. L., & Cornelissen, F. W. (2012). Comparing crowding in human and ideal observers. *Journal of Vision*, *12*(6), 13. https://doi.org/10.1167/12.6.13

Wang, H., He, X., & Legge, G. E. (2014). Effect of pattern complexity on the visual span for Chinese and alphabet characters. *Journal of Vision*, *14*(8), 6. https://doi.org/10.1167/14.8.6

Weiss, Y., Simoncelli, E. P., & Adelson, E. H. (2002). Motion illusions as optimal percepts. *Nature Neuroscience*, *5*(6), 598-604. https://doi.org/10.1038/nn0602-858

Wertheim, T. (1980). Peripheral visual acuity: Th. Wertheim. *American Journal of Optometry and Physiological Optics*, *57*(12), 915-924.

Westheimer, G. (1979). Scaling of visual acuity measurements. *Archives of Ophthalmology*, *97*(2), 327-330.

Whaley, C. P. (1978). Word-Nonword Classification Time. *Journal of Verbal Learning and Verbal Behavior*.

Whitney, C. (2008). Supporting the serial in the SERIOL model. *Language and Cognitive Processes*, *23*(6), 824-865. https://doi.org/10.1080/01690960701828202

Whitney, D., & Levi, D. M. (2011). Visual Crowding: a fundamental limit on conscious perception and object recognition. *Trends in cognitive sciences*, *15*(4), 160-168. https://doi.org/10.1016/j.tics.2011.02.005

Yu, D., Legge, G. E., Wagoner, G., & Chung, S. T. L. (2014). Sensory factors limiting horizontal and vertical visual span for letter recognition. *Journal of Vision*, *14*(9). https://doi.org/10.1167/14.9.23

Appendix 1: Adding letter mislocation errors to letter identity errors

To calculate the decrease in letter recognition rate for the letter recognition profile due to letter mislocations, we used the following method: For a given mislocation coefficient, the distribution probabilities of the three successive letter positions can be easily calculated. It corresponds to three normal probability distributions that we will call $D1$, $D2$, and $D3$ for the three positions of the letters $l1$, $l2$, and $l3$ ($l1 \neq l2 \neq l3$) originally presented at letter slots $s1$, $s2$, and $s3$ (eccentricities $e1$, $e2$, and $e3$). In this case, if the mislocation coefficient is α, $D1$ is a normal distribution centered on $s1$ with a standard deviation $\alpha * e1$, $D2$ is a normal distribution centered on $s2$ with a standard deviation $\alpha * e2$, and $D3$ is a normal distribution centered on $s3$ with a standard deviation $\alpha * e3$.

We can then calculate the probability to perceive the letter $l2$ at any position within the three possible letter slots that would be:

$$P(l2\, perceived\, at\, one\, of\, the\, three\, position\, slots) =$$

$$1-\left(1-\int_{s1}\frac{D2(x)}{D1(x)+D2(x)+D3(x)}dx\right)*\left(1-\int_{s2}\frac{D2(x)}{D1(x)+D2(x)+D3(x)}dx\right)$$

$$*\left(1-\int_{s3}\frac{D2(x)}{D1(x)+D2(x)+D3(x)}dx\right)$$

$1-P(l2\, perceived\, at\, one\, of\, the\, three\, position\, slots)$ is the probability that the letter $l2$ is not an (mislocalization) answer despite a correct identification.

The probability to have $l2$ at the correct position (slot $s2$) would be:

$$P(l2\, perceived\, at\, its\, correct\, position\, slot) = \int_{s2}\frac{D2(x)}{D1(x)+D2(x)+D3(x)}dx$$

Therefore, for each letter recognition rate $P(i)$ from the identity letter recognition profile, we can define the corresponding letter recognition rate $P2(i)$ from the letter recognition profile that takes into account possible mislocations:

$$P2(i) = P(i) * \frac{\int_{s2} \frac{D2(x)}{D1(x) + D2(x) + D3(x)} dx}{1 - \left(1 - \int_{s1} \frac{D2(x)}{D1(x) + D2(x) + D3(x)} dx\right) * \left(1 - \int_{s2} \frac{D2(x)}{D1(x) + D2(x) + D3(x)} dx\right) * \left(1 - \int_{s3} \frac{D2(x)}{D1(x) + D2(x) + D3(x)} dx\right)}$$

Appendix 2: Mislocation coefficient values for Experiment 1a (Trigram presentation) and Experiment 1b (Pentagram presentation)

EXP 1a	LEFT	CENTER	RIGHT
S1	0.10	0.16	0.08
S2	0	0.11	0.04
S3	0.08	0.22	0.16
S4	0.10	0.17	0.08
S5	0.08	0.20	0.13
Average	0.07	0.17	0.10

Exp 1b			
S1		0.37	
S2		0.37	
S3		0.58	
S4		0.45	
S5		0.38	
Average		0.43	

Appendix 3: Gaussian distribution characteristics for Experiment 1a and 1b: Data represents for each subject and each internal position the values for (a) the left standard deviation, (b) the right standard deviation, and (c) the amplitude.

	(a) sd (left)			(b) sd (right)			(c) amplitude		
EXP 1a	**LEFT**	**CENTER**	**RIGHT**	**LEFT**	**CENTER**	**RIGHT**	**LEFT**	**CENTER**	**RIGHT**
S1	9.26	4.95	6.13	9.26	4.95	6.13	1	1	1
S2	11.25	4.03	5.19	6.3	4.81	13.55	1	1	0.97
S3	30	4.77	4.55	8.31	5.11	9.89	1	0.99	1
S4	7.7	3.41	4.05	6.3	3.78	5.76	0.99	0.96	1
S5	9.44	3.46	3.96	4.86	3.87	8.4	1	1	0.99
Average	13.53	4.12	4.78	6.7	4.6	9.74	1	0.99	0.99

Exp 1b	(a) sd (left) **CENTER**	(b) sd (right) **CENTER**	(c) amplitude **CENTER**
S1	3.04	4.22	1
S2	4.06	3.57	0.94
S3	2.6	3.01	0.97
S4	3.45	3.29	0.95
S5	3.5	3.41	0.93
Average	4.39	3.5	0.96

Appendix 4: Transformation of the general confusion matrix

The general confusion matrix CM_{all} is an average confusion matrix that has been defined based on the totality of the letter recognition trials for our five subjects. Its average letter recognition rate is $mean_{all} = 0.44$. To create the new confusion matrix CM_{new} with a different average letter recognition rate $mean_{new}$, we used two weighting methods using $CM_{max} = \begin{pmatrix} 1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 1 \end{pmatrix}$ and $CM_{min} = \begin{pmatrix} 1/26 & \cdots & 1/26 \\ \vdots & \ddots & \vdots \\ 1/26 & \cdots & 1/26 \end{pmatrix}$:

- if $mean_{new} > mean_{all}$, we calculated p (0 < $p < 1$) so that $p * mean_{all} + (1-p) * mean_{max} = mean_{new}$, so $p * 0.44 + (1-p) * 1 = mean_{new}$, and: $p = \frac{(1-mean_{new})}{0.56}$

Then, we calculated $CM_{new} = p * CM_{all} + (1-p) * CM_{max}$

- if $mean_{new} < mean_{all}$, we calculated p (0 < $p < 1$) so that $p * mean_{all} + (1-p) * mean_{min} = mean_{new}$, so $p * 0.44 + (1-p) * \frac{1}{26} = mean_{new}$, and: $p = \frac{(mean_{new} - \frac{1}{26})}{0.40}$

Then we calculated $CM_{new} = p * CM_{all} + (1-p) * CM_{min}$

1 **Appendix 5: Proportion of word length under-estimated, well-estimated, and over-estimated in**

2 **Experiment 2**

Word presentation	**5-letter words**			**7-letter words**			**9-letter words**		
Word length	**under**	**exact**	**over**	**under**	**exact**	**over**	**under**	**exact**	**over**
S1	0.00	0.96	0.04	0.10	0.79	0.11	0.21	0.74	0.05
S2	0.01	0.96	0.03	0.06	0.86	0.08	0.11	0.86	0.03
S3	0.01	0.91	0.08	0.12	0.79	0.09	0.10	0.79	0.11
S4	0.01	0.92	0.07	0.12	0.72	0.16	0.28	0.61	0.11
S5	0.02	0.93	0.05	0.07	0.76	0.16	0.09	0.81	0.09
Average	0.01	0.94	0.05	0.09	0.78	0.12	0.16	0.76	0.08

3

4