KU LEUVEN

FACULTY OF PSYCHOLOGY AND
EDUCATIONAL SCIENCES

LABORATORY OF BIOLOGICAL PSYCHOLOGY
BRAIN AND COGNITION

UCLouvain

INSTITUTE OF NEUROSCIENCE (IoNS) AND
INSTITUTE OF PSYCHOLOGY (IPSY)

Connecting the dots: behavioural, neural, computational models of visual Braille reading

Filippo Cerpelloni

Doctoral thesis offered to obtain the joint degree of
Doctor of Psychology (PhD) at KU Leuven and
Doctor of Biomedical and Pharmaceutical
Sciences (PhD) at UCLouvain

Supervisors:
prof. dr. Hans Op de Beeck
prof. dr. Olivier Collignon

October 2025

Summary

The ability to read profoundly enriches human life. By recognizing single letters, full words, and their arrangement within phrases, we extract meaningful information from written text—an extraordinary skill made possible by the invention and evolution of writing systems. On the neural level, visual reading materializes as a complex network of areas, both visual and linguistic ones, and it involves different stages of the visual system (Dehaene et al., 2005; Yeatman & White, 2021) that culminate in a region of the brain deputed to the recognition of written words, the Visual Word Form Area (VWFA; Cohen et al., 2000, 2002; McCandliss et al., 2003) before being linked to the language network. The exact role of this area, and how it relates visual information to language, has not been completely elucidated. Given that reading commonly relies on vision, VWFA was initially conceptualized as based on the co-option of the visual processing structure that is already in place in sighted individuals (Dehaene et al., 2005; Dehaene & Cohen, 2011). In contrast, cross-modality studies state that the visual word form area may have a different, multifaceted, role in processing linguistic, and symbolic information more in general (Price & Devlin, 2003, 2011). To clarify the nature of the selectivity of VWFA, I studied the visual processing of scripts with a focus on the visual reading of Braille, a script that is not based on line junctions, the basic features that make canonical alphabets. In a first experiment (chapter 2), I investigated the reading acquisition of Braille, and compared it to a similar script based on line junctions (Line Braille). I show how the strength of the mapping between a familiar and a novel orthography drives learning more than the visual features that compose those orthographies. Through neuroimaging experiments

(chapter 3), I recruited expert visual Braille readers to show that VWFA tunes to Braille, similar to how it tunes to a canonical Latin alphabet. Moreover, I present an organization of VWFA where the information of different stimuli is organized according to its linguistic content, in both Braille and Latin alphabet, but in segregated neuronal populations. Lastly, I present results from a computational approach aimed at reproducing reading in Deep Neural Networks (DNNs). I show how the synthetic model fails to reproduce the data and why this is to be attributed to the lack of linguistic knowledge. Finally, I merge the information coming from the separate studies to offer a comprehensive view of VWFA and of the influences of language and vision over reading.

Samenvatting

Het vermogen om te kunnen lezen verrijkt het menselijk leven aanzienlijk. Door het herkennen van afzonderlijke letters, volledige woorden, en hun rangschikking binnen zinnen, wordt zinvolle informatie uit geschreven tekst geëxtraheerd – een unieke vaardigheid mogelijk gemaakt door de uitvinding en ontwikkeling van schriftsystemen. Op neurologisch niveau manifesteert visueel lezen zich in een complex netwerk van zowel visuele als taalkundige gebieden. Het proces van visueel lezen doorloopt eerst verschillende stadia van het visuele systeem (Dehaene et al., 2005; Yeatman & White, 2021). Nadien wordt deze informatie doorgestuurd naar het hersengebied verantwoordelijk voor het herkennen van geschreven woorden, genaamd de Visual Word Form Area (VWFA; Cohen et al., 2000, 2002; McCandliss et al., 2003), om tot slot gekoppeld te worden aan het taalnetwerk. Tot op heden is de precieze rol van de VWFA in visueel lezen en hoe visuele informatie gerelateerd is aan taal nog niet volledig begrepen. Aangezien lezen doorgaans afhankelijk is van het gezichtsvermogen, werd de VWFA aanvankelijk geconceptualiseerd als het resultaat van de co-optatie van bestaande visuele verwerkingsstructuren bij ziende personen (Dehaene et al., 2005; Dehaene & Cohen, 2011). Daarentegen wijzen cross-modale studies erop dat het visuele woordvormgebied, in het algemeen, een andere, veelzijdigere rol kan spelen in de verwerking van linguïstische en symbolische informatie (Price & Devlin, 2003, 2011). Om de werking ervan te verduidelijken, heb ik mij gericht op het observeren van de visuele verwerkingsfase en bijzondere schriften te bestuderen die afwijken van de gebruikelijke gevallen. In een eerste experiment (hoofdstuk 2) onderzocht ik het leren lezen van braille, en

vergeleek ik dit met een soortgelijk schrift op basis van lijnverbindingen (lijn braille). Ik toon aan hoe de sterkte van de koppeling tussen een bekende en een nieuwe orthografie het leerproces meer aanstuurt dan de visuele kenmerken waaruit die orthografieën zijn opgebouwd. Bijkomend, heb ik neuroimaging-experimenten bij visuele braillelezers (hoofdstuk 3) gebruikt om aan te tonen dat de VWFA zich op een vergelijkbare manier afstemt op braille als op het canonieke Latijnse Alfabet. Bovendien presenteer ik een organisatie van de VWFA waarin de informatie van verschillende stimuli wordt georganiseerd op basis van de linguïstische inhoud, zowel in braille als in het Latijnse alfabet, maar in gescheiden neuronale populaties. Ten slotte presenteer ik de resultaten van een computationele benadering gericht op het nabootsen van lezen in diepe neurale netwerken (DNN's). Ik laat zien hoe het artificieel model er niet in slaagt de gegevens te reproduceren en waarom dit te wijten is aan het gebrek aan linguïstische kennis. Aansluitend integreer ik de bevindingen uit de afzonderlijke studies om een omvattend overzicht te bieden van de VWFA en de invloed van taal en visie op lezen.

Resumé

La capacité de lire enrichit profondément la vie humaine. En reconnaissant les lettres individuelles, les mots complets et leur disposition au sein des phrases, nous extrayons des informations significatives du texte écrit — une compétence extraordinaire rendue possible par l'invention et l'évolution des systèmes d'écriture. Au niveau neuronal, la lecture visuelle se matérialise sous la forme d'un réseau complexe de zones, tant visuelles que linguistiques, et implique différentes étapes du système visuel (Dehaene et al., 2005 ; Yeatman & White, 2021) qui aboutissent dans une région du cerveau dédiée à la reconnaissance des mots écrits, l'aire visuelle de la forme des mots (VWFA ; Cohen et al., 2000, 2002 ; McCandliss et al., 2003), avant d'être reliées au réseau linguistique. Le rôle exact de cette aire et la manière dont elle relie les informations visuelles au langage n'ont pas encore été complètement élucidés. Étant donné que la lecture repose généralement sur la vision, la VWFA a d'abord été considérée comme étant basée sur la cooptation de la structure responsable du traitement visuel déjà en place chez les personnes voyantes (Dehaene et al., 2005 ; Dehaene & Cohen, 2011). Cependant, des études intermodales indiquent que la zone de forme visuelle des mots pourrait avoir un rôle différent et multiforme dans le traitement des informations linguistiques et symboliques en général (Price & Devlin, 2003, 2011). Afin de clarifier son fonctionnement, j'ai décidé d'observer l'aire du traitement visuel et d'étudier des écritures particulières qui s'écartent des cas canoniques. Dans une première expérience (chapitre 2), j'ai étudié l'acquisition de la lecture du braille et l'ai comparée à une écriture similaire basée sur des jonctions de lignes (braille linéaire). Je montre comment la force de la correspondance

entre une orthographe familière et une orthographe nouvelle stimule l'apprentissage plus que les caractéristiques visuelles qui composent ces orthographes. À travers des expériences de neuroimagerie (chapitre 3), j'ai recruté des lecteurs experts en braille visuel pour montrer que la VWFA s'adapte au braille, de la même manière qu'elle s'adapte à l'alphabet latin canonique. De plus, je présente une organisation de la VWFA où les informations des différents stimuli sont organisées en fonction de leur contenu linguistique, tant en braille qu'en alphabet latin, mais dans des populations neuronales distinctes. Enfin, je présente les résultats d'une approche computationnelle visant à reproduire la lecture dans les réseaux neuronaux artificiels (DNN). Je montre comment le modèle synthétique ne parvient pas à reproduire les données et pourquoi cela est dû à un manque de connaissances linguistiques. Enfin, je fusionne les informations issues des différentes études afin d'offrir une vue d'ensemble de la VWFA et des influences du langage et de la vision sur la lecture.

Popularized summary

The ability to read is at the basis of modern human life. We recognize single letters, full words, phrases and we extract meaningful information from written text. In the brain, reading involves a complex network of areas, both in the visual system and in the language network (Dehaene et al., 2005; Yeatman & White, 2021). In-between these systems there is an area, the Visual Word Form Area (VWFA; Cohen et al., 2000, 2002; McCandliss et al., 2003), which is most active when people see written words. The exact role of this area has yet to be fully clarified: the visual features that compose the letters are merged into lines and words, similarly for what happens with the regular object we (Dehaene et al., 2005; Dehaene & Cohen, 2011), but the language information of the words also plays a role and leads to an idea of VWFA as a region that binds vision and language more in general (Price & Devlin, 2003, 2011). To fully understand how VWFA works, I studied how people learn and read Braille in the visual modality, visually recognizing the patters of dots. Comparing Braille with Latin alphabets allow us to separate and compare the contributions of vision and language to reading. In a first experiment (chapter 2), I compared how people learn Braille and a custom script where we joined the dots to form lines (Line Braille). Minor differences between alphabets are present, meaning that learning was determined by the linguistic associations and limitedly by visual aspects. In chapter 3, I look at the brain responses and find that, in individuals with many years of experience reading Braille visually, VWFA responds to Braille. On a more refined scale, I show how VWFA responds differently for words and word-like stimuli that differ based on linguistic information. VWFA shows this pattern for Braille too but not in a general way,

keeping the alphabets separated. Lastly, I show how a computerized model of vision, without language information, cannot do the same task as humans, and that this difference is linked to the missing language knowledge. Finally, I combine the results of the three experiments to present a global view of VWFA and of the influences of language and vision over reading.

Populaire samenvatting

Het vermogen om te kunnen lezen vormt de basis van het moderne menselijke leven. We herkennen afzonderlijke letters, volledige woorden en zinnen, om zo zinvolle informatie uit geschreven tekst te halen. In de hersenen omvat lezen een complex netwerk van gebieden, zowel in het visuele systeem als in het taalnetwerk (Dehaene et al., 2005; Yeatman & White, 2021). Tussen deze systemen bevindt zich een gebied, de Visual Word Form Area (VWFA; Cohen et al., 2000, 2002; McCandliss et al., 2003), dat het meest actief is wanneer mensen geschreven woorden zien. De precieze rol van dit gebied is nog niet volledig duidelijk: de visuele kenmerken waaruit de letters bestaan worden samengevoegd tot regels en woorden, net zoals dat gebeurt met gewone objecten (Dehaene et al., 2005; Dehaene & Cohen, 2011), maar de taalkundige informatie van de woorden speelt ook een rol. Dit leidt tot het idee dat de VWFA een gebied is dat visie en taal in het algemeen met elkaar verbindt (Price & Devlin, 2003, 2011). Om volledig te begrijpen hoe de VWFA werkt, heb ik onderzocht hoe mensen braille leren en lezen in de visuele modaliteit, waarbij ze de patronen van punten visueel herkennen. Door braille te vergelijken met het Latijnse alfabet kunnen we de bijdragen van visie en taal aan lezen scheiden en vergelijken. In een eerste experiment (hoofdstuk 2) heb ik vergeleken hoe mensen braille en een aangepast schrift waarbij we de punten samenvoegen tot lijnen (lijnbraille) leren. Er zijn kleine verschillen tussen de alfabetten, wat betekent dat het leren werd bepaald door de taalkundige associaties en in beperkte mate door visuele aspecten. In hoofdstuk 3 kijk ik naar de hersenreacties en ontdek ik dat bij personen met vele jaren ervaring in het visueel lezen van braille, de VWFA reageert op braille. Op een meer

verfijnde schaal laat ik zien hoe de VWFA verschillend reageert op woorden en woordachtige stimuli die verschillen op basis van taalkundige informatie. De VWFA vertoont dit patroon ook voor braille, maar niet op een algemene manier, waarbij de alfabetten gescheiden blijven. Ten slotte laat ik zien hoe een geautomatiseerd model van het gezichtsvermogen, zonder taalkundige informatie, niet dezelfde taak kan uitvoeren als mensen, en dat dit verschil verband houdt met het ontbreken van taalkennis. Tot slot combineer ik de resultaten van de drie experimenten om een globaal beeld te geven van VWFA en van de invloeden van taal en gezichtsvermogen op lezen.

Résumé vulgarisé

La capacité à lire est à la base de la vie humaine moderne. Nous reconnaissons les lettres individuelles, les mots entiers, les phrases et nous extrayons des informations significatives à partir d'un texte écrit. Dans le cerveau, la lecture implique un réseau complexe de zones, tant dans le système visuel que dans le réseau du langage (Dehaene et al., 2005 ; Yeatman & White, 2021). Entre ces deux systèmes se trouve une zone, la zone visuelle de la forme des mots (VWFA ; Cohen et al., 2000, 2002 ; McCandliss et al., 2003), qui est particulièrement active lorsque les gens voient des mots écrits. Le rôle exact de cette zone n'a pas encore été entièrement élucidé : les caractéristiques visuelles qui composent les lettres sont fusionnées en lignes et en mots, comme c'est le cas pour les objets courants (Dehaene et al., 2005 ; Dehaene & Cohen, 2011), mais les informations linguistiques des mots jouent également un rôle et conduisent à considérer la VWFA comme une région qui relie plus généralement la vision et le langage (Price & Devlin, 2003, 2011). Pour bien comprendre le fonctionnement de la VWFA, j'ai étudié la manière dont les gens apprennent et lisent le braille dans la modalité visuelle, en reconnaissant visuellement les motifs de points. La comparaison entre le braille et l'alphabet latin nous permet de séparer et de comparer les contributions de la vision et du langage à la lecture. Dans une première expérience (chapitre 2), j'ai comparé la manière dont les gens apprennent le braille et un script personnalisé dans lequel nous avons relié les points pour former des lignes (braille linéaire). Il existe des différences mineures entre les alphabets, ce qui signifie que l'apprentissage était déterminé par les associations linguistiques et, dans une moindre mesure, par les aspects visuels.

Au chapitre 3, j'ai examiné les réponses cérébrales et j'ai constaté que, chez les personnes ayant de nombreuses années d'expérience dans la lecture visuelle du braille, la VWFA réagit au braille. À une échelle plus fine, je montre comment la VWFA réagit différemment aux mots et aux stimuli de type verbal qui diffèrent en fonction des informations linguistiques. La VWFA présente également ce schéma pour le braille, mais pas de manière générale, en séparant les alphabets. Enfin, je montre comment un modèle informatisé de la vision, sans informations linguistiques, ne peut pas accomplir la même tâche que les humains, et que cette différence est liée à l'absence de connaissances linguistiques. Enfin, je combine les résultats des trois expériences pour présenter une vue d'ensemble de la VWFA et des influences de la langue et de la vision sur la lecture.

Published or accepted journal articles, and/or articles submitted for or worthy of publication

- Chapter 2: Connecting the dots: similar visual orthographic acquisition for Braille or line junctions
  - Authors: Cerpelloni F., Collignon O. *, Op de Beeck, H. *
  - Original title: Connecting the dots: similar visual orthographic acquisition for Braille or line junctions
  - First author: Cerpelloni, F.
  - Publication status: Publicly available as pre-print on *PsyArXiv* and under review in *Psychonomic Bulletin and Reviews*
  - DOI: https://osf.io/preprints/psyarxiv/y3sg2_v1

- Chapter 3: How learning to read visual Braille co-opts the reading brain
  - Authors: Cerpelloni F., Van Audenhaege A., Matuszewski J., Gau R., Battal C., Falagiarda F., Op de Beeck, H. *, Collignon O. *
  - Original title: How learning to read visual Braille co-opts the reading brain
  - First author: Cerpelloni F.
  - Publication status: Accepted for publication in *Journal of Cognitive Neuroscience*
  - DOI: https://doi.org/10.1162/jocn_a_02341

Additional forms of publications, or papers with more than two first authors

- None

Relationship between doctoral thesis and master's theses or other papers

- **Chapter related to the own master's thesis or paper**
  - None

- **Chapter related to another master's thesis or paper**
  - Filippo Cerpelloni was the daily supervisor of two master theses (Emma Gastoud, UCLouvain; Tom Poel, KU Leuven) that collaborated to behavioural testing of individuals learning either Braille or Line Braille (Chapter 2). Emma Gastoud explored a pilot study, in a different language and with a restricted number of participants. Tom Poel was involved in the final experiment, helping in the development of pilot experiments and in the initial phase of the main study. These theses explored some initial and partial aspects, further developed and strongly extended by Filippo Cerpelloni.

Statement about the use of Gen-AI

- Generative AI was not used in the writing of this dissertation.

Table of Contents

Chapter 1. Introduction

Being able to read is a uniquely human ability that involves the recognition of single letters and full words, followed by the extraction of meaningful information. From a linguistic perspective, the basis of reading is interpreting a visual representation (graphemes) of spoken language (phonemes) (Brem et al., 2010; Dębska et al., 2024). In this introduction, I will first present the visual system, how it integrates information to recognise objects in general and in the specific case of written words. On the neural level, visual reading involves a complex network of areas, both visual and linguistic ones, and it includes different stages of visual processing (Dehaene et al., 2005; Yeatman & White, 2021) that culminate in a region of the brain deputed to the recognition of written words, the Visual Word Form Area (VWFA; Cohen et al., 2000, 2002; McCandliss et al., 2003), before it accesses the language network. Moreover, for how useful it can be, reading is not innate but is instead a consequence of extensive training (Dehaene et al., 2010; Dehaene-Lambertz et al., 2018; Van Atteveldt et al., 2004). Numerous studies investigated reading acquisition from different perspectives, in humans (Baker et al., 2007; Cerpelloni et al., 2024; Martin et al., 2019; Moore et al., 2014) and primates (Rajalingham et al., 2020). The exact role of VWFA has not been completely elucidated. In this chapter, we will then explore the literature and discuss how VWFA relates visual information to language, with a focus on Braille and what kind information we obtain from studying it in different sensory modalities. Lastly, I will present the current methodologies I implemented to study the brain. All these aspects will serve to contextualize the following chapters, in which I will present, and later

discuss, my work on the organisation of the human visual system and its specialisation towards reading.

The visual system

Visual information from the outside environment flows into the brain through a series of processing stages, from the electrical response of photoreceptors on the retina, to the Lateral Geniculate Nucleus (LGN) of the thalamus, to the primary visual cortex (V1) in the occipital lobe of the brain (Figure 1) (Bear et al., 2006). Within V1 reside neurons that have shown preferential responses, in terms of maximal activation, for the direction or for the orientation of a presented stimulus (Hubel & Wiesel, 1962). From this simple tuning to oriented bars, the full range of visual processing is achieved through connections spanning from V1 and V2 to several other cortical regions.

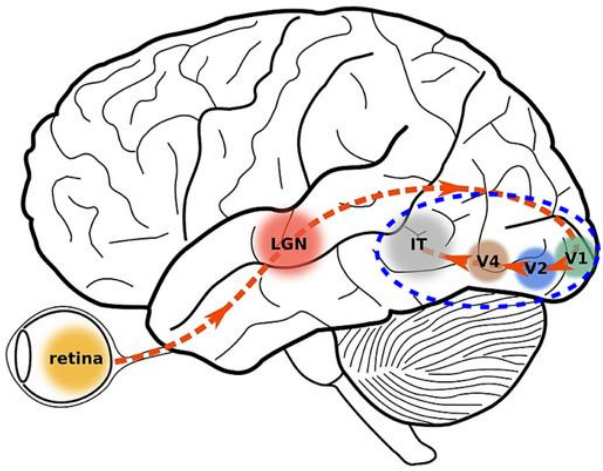

Figure 1 – Schematic representation of the early stages of visual processing and the ventral visual stream. Visual information from the retinas is first processed in the Lateral Geniculate Nucleus (LGN) and then transferred to the primary visual cortex (V1). The ventral visual stream, involved in object recognition, spans from V1-V2 to V4 and successively to the Interior Temporal cortex (IT), where different specialized areas are present, each responding to a stimulus category (see Figure 2). Adapted from: Kubilius, Jonas (2017). Ventral visual stream. doi.org/10.6084/m9.figshare.106794.v3

Two main streams processing different aspects of the environment flow from the early visual cortex. The dorsal stream projects to areas that elaborate object motion (V5/MT) and complex movement (MST), and processes the spatial information of the objects we see – often referred to as the "where" pathway. In the ventral stream – commonly known as the "what" pathway (Figure 1), the perceived lines composing the objects we see are progressively integrated into full objects. Visual information coming from the early visual cortex is mainly elaborated in V4, an area important for colour and shape perception, and transferred to the Inferior Temporal cortex (IT), where different patches of cortex show sensitivity to different categories of visual

stimuli (Bear et al., 2006). In these regions stimuli are processed for their identity and features in a process known as visual object recognition (Biederman, 1987; Logothetis & Sheinberg, 1996; Riesenhuber & Poggio, 1999, 2002). The best-known example of object recognition present in the literature is the recognition of faces (Kanwisher et al., 1997). Given their importance for social communications, processing faces and their emotional value is a crucial skill for survival. Different patches of the ventral occipito-temporal cortex (VOTC) have been found to be tuned to processed them, to the point that they also respond to face-like configurations, producing the effect of pareidolia (Liu et al., 2014).

Functional categorization of objects

From a functional perspective, the Inferior Temporal cortex (IT), also known as the Ventral Occipito-Temporal Cortex (VOTC), hosts a mosaic of specialized areas (Figure 2), each showing a preference for particular stimuli or configurations (Grill-Spector & Malach, 2004; Grill-Spector & Weiner, 2014; Op de Beeck et al., 2019; For a review: Bracci & Op De Beeck, 2023). Within VOTC we can find an arrangement of such areas according to different principles of eccentricity (Hasson et al., 2002), animacy (i.e. animate categories like animals being represented more laterally) (Connolly et al., 2012; Yargholi & Op De Beeck, 2023). The shape of the objects also plays a role, as the lateral occipital cortex (LOC) has been shown to integrate the shape information from earlier processing stages into full objects (Cichy et al., 2011; Kim et al., 2009; Kourtzi, 2003; Kourtzi & Kanwisher, 2000, 2001; Op de Beeck et al., 2008). Ultimately,

we can identify regions, other the aforementioned areas for faces (Kanwisher et al., 1996, 1997), that preferentially respond to categories such as body parts (Downing et al., 2001), scenes (Epstein et al., 1999; Epstein & Kanwisher, 1998), words (Cohen et al., 2000, 2002) and objects (Malach et al., 1995).

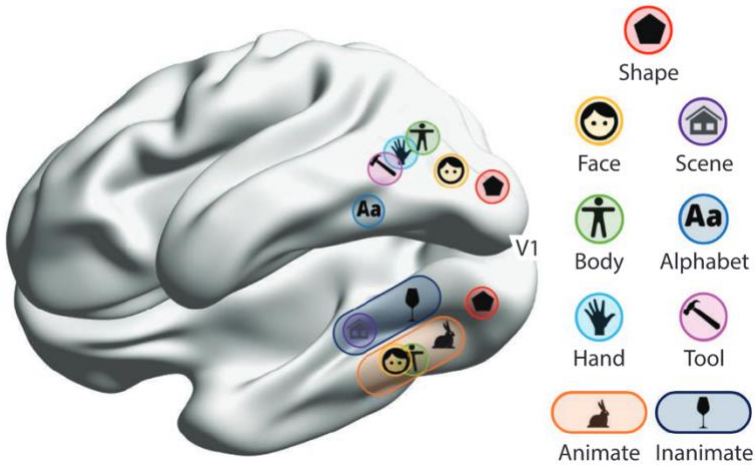

Figure 2 – functional organization of the ventral occipito-temporal cortex. The ventral occipitotemporal cortex (VOTC) hosts several areas selective for shape processing, faces, body, hands, scenes, tools, words. These areas are clustered according to principles of eccentricity (distance on the retina from the fovea) and along an animate-inanimate axis. Adapted from Bracci & Op De Beeck, 2023.

These categories are not randomly assigned but are instead relevant for human behaviour. As a notable example, faces and facial expressions are crucial for our social interactions: misjudging an emotional facial expression can determine a different outcome toward a given person. A similar point can be made for body parts, scenes (recognising potential threats or sources of food in our environment), objects (elements we can interact with). Words, contrarily to other types of stimuli benefitting from specialized processing, are not innate but a recent human invention and yet show similar preferential

areas in the same cortical region. For this reason, words are thought to build upon an existing system of progressively specialized receptive fields (Dehaene et al., 2005; Dehaene & Cohen, 2007), culminating in VOTC. Additionally, literacy acquisition is necessary for the emergence of a dedicated area (Dehaene et al., 2015; Dehaene-Lambertz et al., 2018). The systematic localization of word-selective processing in VOTC, and the requirement of training for it to emerge, make words and reading key elements to study the principles of object recognition and its neural correlates.

The functional profile of the Visual Word Form Area

Written words are an object category that carries two levels of information. In addition to the visual features of the letters, words possess various linguistic (i.e. semantic, phonological, orthographic) contents, much more explicit than other categories. This duality in the levels of information makes words a useful model to study how low- and high-level properties interact and are integrated in visual processing. Additionally, the expertise needed for a specialized area to emerge (Dehaene et al., 2015), the Visual Word Form Area (VWFA) (Cohen et al., 2000, 2002), is of particular interest for studying how the reading network changes as a consequence of experience-induced plasticity and how, more in general, the brain adapts to process novel stimuli.

The first studies that identified VWFA defined its function as the processing of the orthographic content (i.e. form) of words (Dehaene & Cohen, 2007, 2011; Li et al., 2024). This orthographic tuning hypothesis builds on the presence of

an innate visual stream with receptive fields that are specialized for progressively more complex stimuli (Logothetis & Sheinberg, 1996), and states that this system is co-opted when reading to process written words (Dehaene et al., 2005; Dehaene & Cohen, 2007), as shown in Figure 3. In this view, VWFA is the node where lines, junctions, and bigrams are combined to form word and pseudowords representations (Dehaene et al., 2005; Dehaene & Cohen, 2011). This theory found support in studies that showed a preference for line junctions in object recognition and in word recognition (Szwed et al., 2009, 2011), with neural correlates supporting this preference (Szwed et al., 2011), and in neuroimaging experiments showing, in the visual stream, a posterior-to-anterior gradient of preferential response for full words compared to lines, letter, bigrams (Vinckier et al., 2007). Additionally, other studies showed that known and foreign characters produced similar VWFA activations (Vogel et al., 2012; Xue & Poldrack, 2007), supporting the idea that VWFA processes the shape of the stimuli, probably in a general manner rather than being specialized for scripts (Roberts et al., 2013).

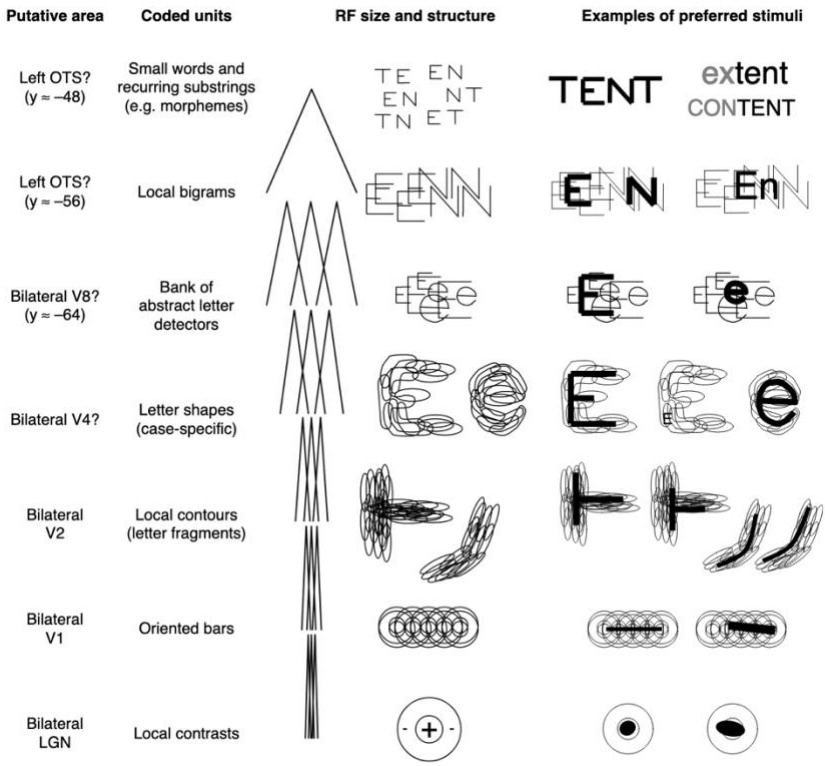

Figure 3 – Orthographic tuning hypothesis. The model of invariant word recognition states that the increasing complexity of the receptive fields of the visual hierarchy is co-opted to process orthographic infomration. In this view, oriented bars perceived in V1 are integrated in contours, letter shapes, small words across the ventral visual stream. Image from Dehaene et al., 2005.

Challenging the orthographic tuning hypothesis, a different view is that VWFA engages not only in the orthographic representations of words, but rather in multiple tasks involving language (Price & Devlin, 2003, 2011). Many studies focused on the different types of visual stimuli that could elicit a preferential response in VWFA and looked, for example, at the acquisition of novel scripts like Hebrew (Baker et al., 2007), Korean Hangul (Xue & Poldrack, 2007), or custom alphabets. Interestingly, these findings extend beyond typical orthographic scripts. Other studies took existing categories in VOTC, faces

(Moore et al., 2014) or houses (Martin et al., 2019), and trained participants to link them to phonemes, showing VWFA responses once the stimuli were mapped onto linguistic properties. All these studies suggest that VWFA can represent linguistic information for a wide array of visual stimuli, as long as a link is made between visual and linguistic properties. However, a counter-argument could be made about these studies, as the scripts involved, for how distant they are from canonical scripts (Changizi et al., 2006), can be framed as line-based scripts and therefore being subject to the same integration of low-level visual features. Thus, whether a logographic alphabet, a house, or a face is presented, it is composed of the same elements that are thought to be co-opted by the reading network.

Task modulation provides additional evidence over the role of VWFA. Looking at a finer scale of neural activations reveals that the relationships between words are based on what is asked to the experiment participants, whether to associate words based on thematic (same environment, e.g. 'teacher'-'classroom') or taxonomic (people, object, locations; e.g. associating 'teacher' to 'doctor') categorization (X. Wang et al., 2018). Additionally, general engagement in a task results in higher VWFA activation (White et al., 2023). Such results highlight how the role of VWFA is not limited to the processing of the orthographic information (invariant of the associations between words) but varies with the different forms of top-down modulating signals from other parts of the language network. Moving away from static stimuli, a recent study used a multivariate approach to explore the visual processing of lip reading (Van Audenhaege et al., 2024). The authors first localized VWFA for written words, as well as the fusiform face area (FFA) and the extrastriate body area

(EBA), and then asked the participants to see a series of phonemes and lip movements representing both visemes, the visual side of the movements necessary to produce a phoneme (i.e. phonological information) and movements that were not associated to linguistic information. Multivariate analyses showed that in VWFA it is possible to decode both phonemes and silent lip movements. In parallel, the authors could not cross-decode the phonological information across senses (i.e. across visemes and phonemes), hinting at representations that are not integrated in an amodal manner (Van Audenhaege et al., 2024).

Studies showing VWFA responses for linguistic information coming from different sensory modalities further support the view of VWFA interconnected with the language network. While Braille will be explored more in depth in later paragraphs, it is worth noting already that VWFA has been involved in tactile Braille reading in both blind (Büchel et al., 1998; Reich et al., 2011) and sighted individuals (Siuda-Krzywicka et al., 2016). More important in this section are studies on speech perception, where VWFA processing for spoken words have been shown in addition to written ones (Burton et al., 2003; Dębska et al., 2019; Dzięgiel-Fivet et al., 2021; Pattamadilok et al., 2019; Planton et al., 2019; S. Wang et al., 2022). Pattamadilok and colleagues, in particular, developed a TMS paradigm to distinguish between different hypotheses about the presence of VWFA activation for spoken words and concluded that neural populations coding specifically for audition co-exists with neural populations dedicated to written words (Pattamadilok et al., 2019). This result is in line with the lack of cross-modal decoding in Van Audenhaege

and colleagues (2024) and may underly a more fragmented view of VWFA, with different populations of neurons specific for scripts or modalities.

Further evidence about the interaction between VWFA and the language network comes from studying the connectivity profile of VWFA. Indeed, anatomical connections between VWFA and areas of the language network are found in adults (Bouhali et al., 2014; Yeatman et al., 2013), indicating a link between the processing in VWFA and a more linguistic processing. Saygin and colleagues, studying children at five and eight years old (before and after learning to read in school) and by defining VWFA when children acquired reading skills, found that the localization of VWFA in eight years old children could be predicted by the connectivity profile of VOTC at five years old (Saygin et al., 2016). Later, this study was expanded by Li and colleagues, who showed a strong connectivity profile between language regions and VWFA even in infants (Li et al., 2020). Altogether, these studies show that, while reading acquisition is a requirement, the location and the functional profile of VWFA are intertwined, and might even be determined, by language. Similarly, Wang and colleagues approached the investigation of VWFA's functional role using graph theory. The authors compared fMRI data from participants listening to clear or degraded speech signal, and constructed a network of regions of interest. Complementing previous studies, they showed how the connectivity in left-vOT is modulated by top-down and bottom-up information and concluded that the functional profile of left-vOT depends on the stimulus, on the context, and on the demands of the task (S. Wang et al., 2022).

Recent developments in analyses and in high-resolution fMRI made possible to investigate the organisation of VOTC on a deeper, fine-grained scale. This approach unravelled how many of the areas initially conceived as one can be in fact divided in two or more sub-areas, each with a particular preference within the same category (for a review, Grill-Spector & Weiner, 2014). In line with other stimulus-specific patches in VOTC, different studies highlight a more fine-grained profile of VWFA. Lerma-Usabiaga and colleagues used a multimodal approach to show that two different areas that respond to written words reside within VOTC: the middle Occipito-Temporal Sulcus (mOTS), activated mostly by top-down information flow and sensitive to lexical information, and the posterior Occipito-Temporal Sulcus (pOTS), showing preference for bottom-up perceptual signals (Lerma-Usabiaga et al., 2018). These two sub-areas seem to reflect two aspects of reading: the visual perception of words and the connections to language areas. Zhan and colleagues identified not two but multiple script-specific patches. Studying English-Chinese bilinguals, the authors found clusters that responded more to Chinese than English around the left fusiform gyrus, while in English-French bilinguals (two scripts that build on the same alphabet), the patches identified for words responded to both scripts (Zhan et al., 2023). Pillet and colleagues mapped VOTC on a more detailed scale by presenting visual stimuli from multiple categories to their participants (Pillet et al. 2024a; Pillet et al., 2024b). Regarding VWFA, they identified three distinct clusters of activation across individuals, of which one area previously attributed to letter sensitivity. Overall, new technological and methodological advancements are enabling new understandings into VWFA, VOTC, and the whole brain. More specifically, the identification of two, or multiple, sub-areas of VWFA with different roles,

opens to the possibility of the two main hypotheses co-existing in different subparts of the same area. In this case, mapping the contributions of visual and linguistic properties could help elucidate the processing of words.

Overall, the literature on VWFA offers a multifaceted picture about the location, the stimuli processed, and the influence of visual and linguistic properties. Two important reviews (Yeatman & White, 2021; Dębska et al., 2023) summarize the key findings about the biological and cognitive aspects of VWFA. A consensus can be found on three major points. First, VWFA responds to printed words and pseudowords, in several alphabets. Second, fine spatial scale localization highlights different sub-regions that may underlie different preferences. Third, VWFA responds to words that are not only written, but also presented in widely different modalities, like spoken language (Pattamadilok et al., 2019; Planton et al., 2019), lip movements (Van Audenhaege et al., 2024), touch (Büchel et al., 1998; Reich et al., 2011; Siuda-Krzywicka et al., 2016). These findings question the necessity of the basic visual features required for orthographic processing (i.e. line junctions) and call instead for a more interactive account of VWFA. What remains to be studied is how crucial line junctions are for the specialization of VWFA and what is the interplay between visual and linguistic properties.

Tactile and visual reading of Braille

One approach to elucidate the questions of VWFA's preference for visual combinations of line junctions and of the role played by linguistic information is to study Braille reading. Braille is a script developed for tactile processing, therefore its study mainly involves visually impaired individuals and a different sensory modality than visual reading. This framework allows to investigate how lexical properties are accessed without depending on visual information. In this context, findings from the processing of letters and words can provide relevant information on the impact of linguistic information. De Heering and colleagues found evidence of breaking the mirror invariance (the effect of invariant perception of an object if seen from its left or rights side; effective for object recognition but actively suppressed in reading to recognize similar symbols, like 'b' and 'd') for tactile Braille letters (de Heering et al., 2018), similarly to what has been shown for Latin letters in sighted (Kolinsky et al., 2011; Pegado et al., 2011, 2014) and even in blind individuals (Korczyk et al., 2024). On the neural level, blind individuals activate VWFA when reading tactile Braille words (Büchel et al., 1998; Reich et al., 2011), and the same was found for sighted individuals (Siuda-Krzywicka et al., 2016). Haupt and colleagues explored the spatial dynamics of tactile reading of Braille letters (Haupt et al., 2024). The authors identified multiple areas processing the perceptual representations (defined as more abstract than sensory representations) of Braille letters. Importantly, among somatosensory areas and the early visual cortex (already extensively reported in the literature by Büchel et al., 1998; Dzięgiel-Fivet et al., 2021; Ptito et al., 2008; Sadato et al., 1998, 2004), they found VWFA and LOC and suggested a modality-

independent organization for VWFA and a “hinge” function, binding sensory and perceptual representations, for LOC. Overall, these findings already highlight tactile Braille processing as similar, albeit with longer latency, to canonical visual scripts.

Braille was created for reading in absence of vision. For this reason, its structure diverges from most scripts (Changizi et al., 2006) and does not contain line junctions, but rather relies on the organization of elevated dots on a six-points grid. Figure 4 illustrates the difference between the Latin alphabet (top rows) and Braille (bottom rows). The difference between basic features and the letters similarities are evident. This peculiarity makes Braille a model candidate to test whether such visual properties are necessary to build a linguistic representational structure of readable stimuli and how similar the representations would be.

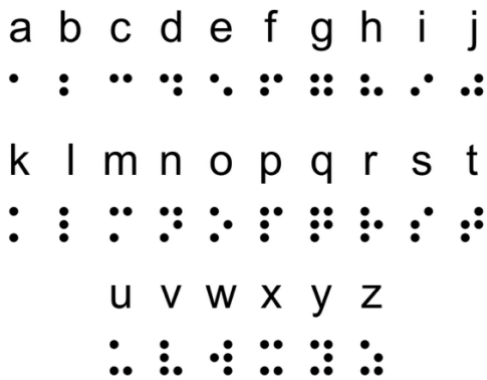

Figure 4 – Braille alphabet. Latin alphabet (top rows) and Braille (bottom rows). Being developed for reading in absence of vision, the Braille characters are not derived from the evolution of strokes but are arbirtrarly assigned in a grid. The patterns of dots representing the first ten letters (‘a’ to ‘j’) is repeated for the following ones with the addition of a dot at the bottom left (position 3). The same structure is then repeated for the last letters, adding an additional dot in the bottom-right position (position 6). The script, developed by Louis Braille, was invented from French, where no letter ‘w’ was used. For this reason, the Braille character for ‘w’ deviates from this structure.

Multiple studies have looked at how VWFA elaborates linguistic input coming from different scripts (Baker et al., 2007; Xue & Poldrack, 2007; Zhan et al., 2023), from other sensory modalities (Pattamadilok et al., 2019; Planton et al., 2019), or from other peculiar scripts and visual forms (Dalski et al., 2024; Van Audenhaege et al., 2024). However, the basic features that make Braille also make it stand out as the minimal substitution of letters maintains the language access the same of the Latin alphabet but is associated with vast changes in the low-level features. In the beginning, Braille was developed as a letter-by-letter mapping from characters of the French alphabet (Braille grade 1). Currently, although improved versions exist (e.g. abbreviations for certain frequent words to ease and speed-up reading; Braille grade 2), such mapping is maintained, at least in scripts based on the Latin alphabet (e.g. English, Dutch, French, Italian). Moreover, among sighted individuals who work with Braille, some are able to read it and do so visually, using predominantly Braille grade 1 and reading letter by letter. Therefore visual reading of Braille relies on the same sensory modality and presents different visual features yet preserving the lexical components of the words: bigrams, phonemes, graphemes, are present in both scripts with the same linguistic representation, varying only in the visual features that compose them. For this reason, visual reading of Braille offers a unique way to distinguish between the contributions of the preserved linguistic content and the different visual properties of words.

Likely due to the rarity of visual Braille reading, only a small set of studies is available in the literature. De Heering and colleagues tested mirror invariance for visual Braille in a group of sighted expert Braille readers and found similar

effects to what they previously observed for tactile Braille in blinds (de Heering & Kolinsky, 2019). Bola and colleagues (Bola et al., 2017) compared script acquisition in two groups learning either visual Braille or Russian (i.e. Cyrillic alphabet) and observed higher accuracy and faster response times for the Cyrillic script compared to Braille. While a broader analysis of this study's limitations can be found in *chapter 2*, it is important to note already that the training paradigms differed between scripts, and that the learning of Braille was both tactile and visual. On the neural level, VWFA responds to visually presented Braille words in sighted individuals trained on a combination of tactile and visual Braille (Gaca et al., 2025; Siuda-Krzywicka et al., 2016). In particular, Gaca and colleagues explored the temporal evolution of tactile Braille learning. Over the course of seven months, participants were asked to learn tactile Braille letters of progressive difficulty. Throughout the training, they performed tactile and visual tests in four fMRI scanning sessions (plus two pre-training and one post-training sessions). The authors reported VWFA activation as early as seven days into the training, highlighting the high plasticity of VWFA when learning a new script. Moreover, they also reported higher activations for visually presented Braille, a script not explicitly learnt but to which participants were nonetheless exposed (Gaca et al., 2025).

Overall, the mechanisms of Braille acquisition seem to parallel those for any other visual script. The activation of VWFA (Büchel et al., 1998; Reich et al., 2011; Siuda-Krzywicka et al., 2016) and the acquisition trajectory (Gaca et al., 2025) resemble the activity and acquisition of common scripts (Dehaene-Lambertz et al., 2018). However, these studies have limitations, mostly in the way Braille is presented to the participants. The lack of differentiation between

tactile and visual Braille limits the range of conclusions that can be drawn about brain plasticity and scripts processing. Moreover, these studies looked at the presence or absence of response and did not explore how Braille is represented within the activations they found. This question is still open and has not been addressed yet by the literature. Multivariate analyses found an effect of the task-modulation in VWFA (X. Wang et al., 2018), meaning that words were represented differently based on the linguistic context. Based on this, it is also possible that words in different languages but with converging meaning are represented in a similar fashion (e.g. 'fox' in English, 'renard' in French). On this line, it becomes then possible that two scripts (Latin alphabet and Braille), mapping into the same linguistic representation (same semantic meaning, same graphemes), are represented in the same manner inside VWFA.

Computational models of vision and reading

The recent development of Deep Neural Networks (DDNs) presents an opportunity to investigate object categorization, and reading, through computerized models of the visual system that complement typical behavioural and neuroimaging methods (see Figure 5 for an example of visual model applied to behaviour and neural responses), but limit the processes co-occurring in the brain. DNNs are models that consist of multiple layers, each one composed of many units performing operations, and implement a function to its input (either the stimulus or the output of a previous layer) with the goal of recognizing patterns in the data. In this view, numerous studies drew parallels between the processing of a Deep Neural Network and the one

of the hierarchy in primate and human vision (Cadieu et al., 2014; Cichy, Khosla, et al., 2016; Cichy, Pantazis, et al., 2016; Khaligh-Razavi et al., 2018; Seeliger et al., 2018; Yamins et al., 2014). Computational models pose therefore many advantages to the field of visual perception, among which the possibility to visualize and isolate the activation state of a layer (i.e. a processing stage).

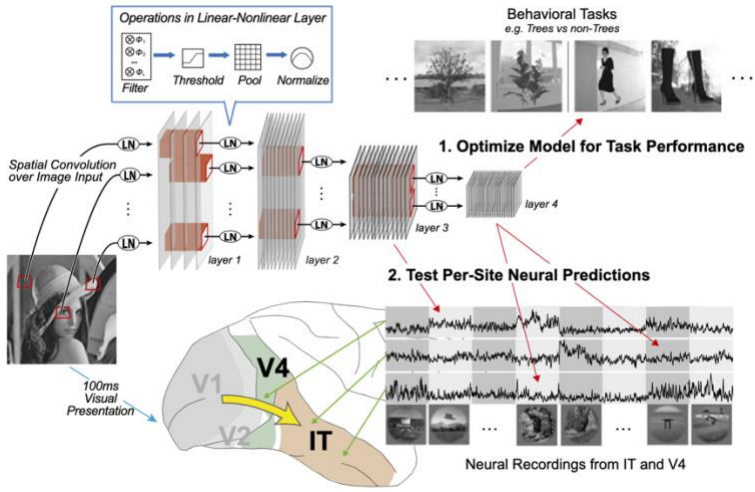

Figure 5 – Example of computational model applied to behavioural and neural data. A hierarchical convolutional neural network is presented with images used to optimize the model. Output from layers of the network is compared with electrophysiology data from microelectrodes implanted in IT. Adapted from Yamins et al. 2014

Recent studies applied Deep Neural Networks to the recognition of letters and words. Testolin and colleagues developed a network of five layers (described as LGN, V1, V2, V4, OTS). The authors trained the network to process images of naturalistic scenes and tested it on the letter-recognition performance (Testolin et al., 2017), supporting the hypothesis that the visual features of common scripts are recycled from the line edges of the objects we see (Bola et al., 2017; Dehaene et al., 2005). Similarly, Janini and colleagues used

AlexNet (Krizhevsky et al., 2017), a benchmark Convolutional Neural Network (CNN), to explore the degree of specialization necessary for letter recognition. The authors compared instances of AlexNet trained either on letters or on ImageNet, a dataset of common objects (Russakovsky et al., 2015), and concluded that the visual features of common objects predict the identification of letters in a similar manner to what happens in human behaviour (Janini et al., 2022). Other researchers resolved to model the different stages of information processing that happen in the brain during word reading (Agrawal & Dehaene, 2024; Hannagan et al., 2021). By using different models of CORnet (Kubilius et al., 2018), a network considered biologically plausible (i.e. with a processing hierarchy similar to the ventral pathway of the visual stream), they first trained the networks to recognize objects and later trained them on the recognition of words alongside objects, simulating the type of learning that children undergo when they start school. After training, the authors observed units selective to letters and ordinal positions in the layer corresponding to Inferior Temporal cortex (IT) (Agrawal & Dehaene, 2024). Overall, these recent studies expand the methodologies available to the investigate reading and offer insights on the purely visual processing of words, separated from interactions with the language network. The way in which the statistical regularities of the words are represented in DNNs remains unexplored, but it can potentially reveal important information over the interplay between visual and linguistic properties.

Methodological notes: Multi-Variate Pattern Analysis (MVPA) and Representational Similarity Analysis (RSA)

Through functional Magnetic Resonance Imaging (fMRI), it is possible to measure the brain activity in response to different kinds of stimulation. In particular, we can visualize which areas are or are not active by mapping the brain into small units, voxels (volume pixels), containing information about a specific piece of the brain. We used two main techniques in this thesis, univariate localization and Multivariate Pattern Analysis (MVPA). Univariate localization consists in the contrast of activations relative to two (or more) categories of stimuli (e.g. intact and scrambled words), to localize which brain areas show a selective response for the property that varies between the categories (e.g. the word's orthography). This approach considers all the voxels responding to a contrast, and spatially in close proximity, as one area, with no major distinctions beside the amount of response. Multi-Variate Pattern Analysis (MVPA) allows for a finer look at the pattern of activity coming from a source, for example the voxels contained in a region of interest (ROI) of the brain, or the activation coming from layers of a given Deep Neural Network (DNN). The pattern of activity in response to one stimulus or class of stimuli is used to train a classifier, an algorithm that can discern the data relative to two or more conditions and produces a prediction about the data. By training over multiple presentations of the data, always excluding a different portion to test the accuracy of the predictions, we obtain a measure of how accurately the classifier can distinguish the input data in different classes. The overall accuracy of the classifier, averaged across iterations of the algorithm, will indicate how well the classifier learnt the differences between data. The classifier accuracy can be interpreted as an index of how differently the ROI or network layer represents stimuli or conditions, which in turn can reflect differences in the stimuli themselves. In this dissertation, I used MVPA

(sometimes referred to as decoding) in *chapter 3*. Specifically, I used a Support Vector Machine (SVM) algorithm to classify the neural patterns of Latin alphabet and Braille stimuli and I used a Leave-One-Run-Out (LORO) approach, meaning that one out of the six functional runs of fMRI data obtained for each participant and script was left out as test at each iteration.

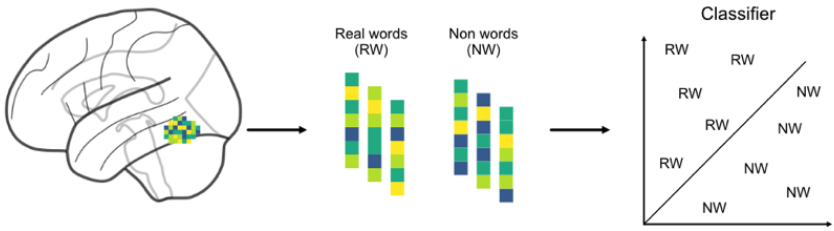

Figure 6 – Multivariate Pattern Analsysis (MVPA). The pattern of voxels related to a stimulus or category is extracted from one Region of Interest (ROI) and used to train a classifier algorithm on the distinction between the patterns. The classifier, tested on patterns previously unseen, produces a prediction (decoding accuracy) based on the distinctions between patterns. High decoding accuracy indicates an easy distinction, meaning substantial differences in the patterns of voxels. Chance-level dceoding indicates instead a difficulty in making an accurate prediction, which underlies a high similarity of the neural patterns, and therefore of the neural processing. In the example, information from VWFA is extracted in relation to the processing of Real words and Non words (strings of consonants).

Representational Similarity Analysis (RSA) is a framework that allows to relate different branches of neuroscientific research (Kriegeskorte, Mur, & Bandettini, 2008; Kriegeskorte, Mur, Ruff, et al., 2008). It relies on a Representational Dissimilarity Matrix (RDM), a matrix that displays the differences between stimuli or classes of stimuli coming from one source (e.g. behavioural data, neuroimaging analyses, Deep Neural Network layer activations, electrophysiology) in an abstract space of dissimilarity. Dissimilarity matrices from multiple sources, modalities, areas can then be compared with one another to understand the relationships between them and between the sources from which they are derived. Figure 6 illustrates an example of the

comparisons that can be done through RSA, for example between different subjects, or different brain areas within one subject. RDMs can be computed through different metrics. Specifically, in this dissertation I will use the MVP classification of brain patterns (chapter 3) and two different distance measures (correlation-based and Euclidean) between the activations across layers of two Deep Neural Network architectures (*chapter 4*).

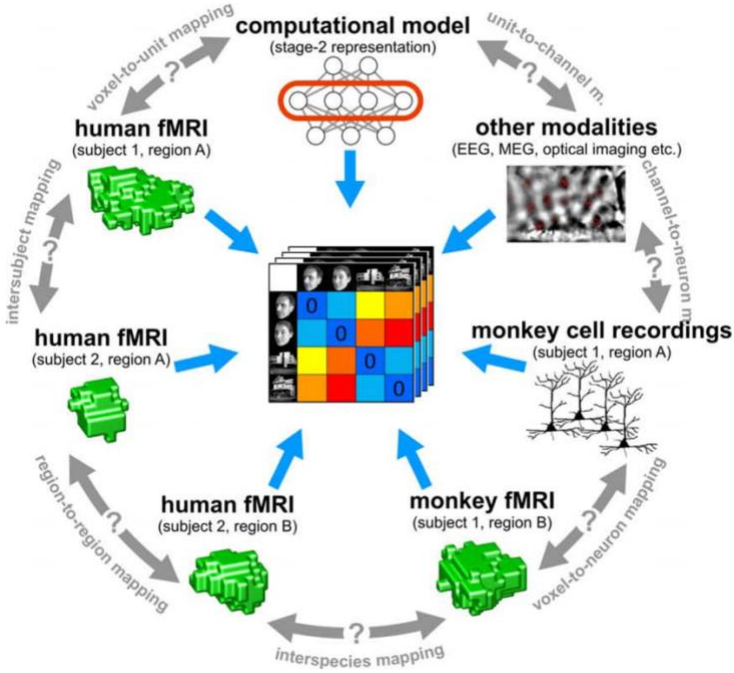

Figure 7 – Representational Similarity Analysis (RSA). Information across different sources (subjects, brain regions, imaging methods) of data can be compared by computing the distances across stimuli within one source. By considering only the relationships between the stimuli, it is possible to compute the (dis)similarities between sources. Adapted from Kriegeskorte et al. (2008)

Overview of the dissertation

In this thesis, I will argue that the neural mechanisms of reading are not yet fully mapped, particularly regarding the contributions of the different words properties and the role they have in shaping the activations of the Visual Word Form Area (VWFA). While reading has traditionally been studied in the context of line-based scripts, the VWFA's adaptability to alternative systems, such as visually processed Braille, remains underexplored. By combining behavioral measurements, neuroimaging analyses, and computational simulations, I investigate how the brain -and specifically VWFA- adapts to process Braille through vision, challenging canonical notions of what is considered necessary for reading (Bola et al., 2017; Changizi et al., 2006).

In *chapter 2*, I present a behavioural study testing the necessity of line junctions for reading. By comparing visual reading acquisition for two unknown scripts, Braille and a line-based counterpart, I demonstrate that comparable levels of recognition can be achieved regardless of the presence or absence of line junctions in the learnt script, albeit with a slight delay in the learning curve when lines are absent.

In *chapter 3*, I build upon this investigation by examining the neural organization of visually processed Braille in individuals with extensive experience reading Braille visually. Using univariate and multivariate fMRI analyses, I show that the functional organization of VWFA exhibits similarities between Braille and Latin-based French. Participants with the ability to process both scripts demonstrate representational similarities (1) extending beyond the ventral occipito-temporal cortex (VOTC) and VWFA to include early visual areas and LOC, and (2) reflecting script-specific characteristics rather than a unified, script-independent representation.

In *chapter 4*, I employ computational simulations to model the visual neural processes underlying reading. These simulations investigate changes in the visual system during both early script acquisition (chapter 2) and extensive exposure to Braille (chapter 3). I show that computational representations do not converge on the same patterns of neural processing and that is a consequence of the limits of a pure visual processing that lacks language representations.

Finally, in *chapter 5*, I synthesize the findings from the individual studies to provide a comprehensive perspective on the behavioral, neural, and computational processes involved in visual Braille reading. Combining multiple modern complementary analytical approaches, this integrated view expands our understanding of VWFA as a mediator between visual input and linguistic aspects of reading, offering insights into the flexibility of neural mechanisms that support this uniquely human skill.

References

Agrawal, A., & Dehaene, S. (2024). Cracking the neural code for word recognition in convolutional neural networks. *PLOS Computational Biology*, *20*(9), e1012430. https://doi.org/10.1371/journal.pcbi.1012430

Baker, C. I., Liu, J., Wald, L. L., Kwong, K. K., Benner, T., & Kanwisher, N. (2007). Visual word processing and experiential origins of functional selectivity in human extrastriate cortex. *Proceedings of the National Academy of Sciences*, *104*(21), 9087–9092. https://doi.org/10.1073/pnas.0703300104

Bear, M. F., Connors, B. W., & Paradiso, M. A. (2006). *NEUROSCIENCE: Exploring the Brain, Fourth Edition*.

Biederman, I. (1987). *Recognition-by-Components: A Theory of Human Image Understanding*.

Bola, Ł., Radziun, D., Siuda-Krzywicka, K., Sowa, J. E., Paplińska, M., Sumera, E., & Szwed, M. (2017). Universal Visual Features Might Be Necessary for Fluent Reading. A Longitudinal Study of Visual Reading in Braille and Cyrillic Alphabets. *Frontiers in Psychology*, *8*. https://doi.org/10.3389/fpsyg.2017.00514

Bouhali, F., Thiebaut de Schotten, M., Pinel, P., Poupon, C., Mangin, J.-F., Dehaene, S., & Cohen, L. (2014). Anatomical Connections of the Visual Word Form Area. *Journal of Neuroscience*, *34*(46), 15402–15414. https://doi.org/10.1523/JNEUROSCI.4918-13.2014

Bracci, S., & Op De Beeck, H. P. (2023). Understanding Human Object Vision: A Picture Is Worth a Thousand Representations. *Annual Review of Psychology*, *74*(1), 113–135. https://doi.org/10.1146/annurev-psych-032720-041031

Brem, S., Bach, S., Kucian, K., Kujala, J. V., Guttorm, T. K., Martin, E., Lyytinen, H., Brandeis, D., & Richardson, U. (2010). Brain sensitivity to print emerges when children learn letter–speech sound correspondences. *Proceedings of the National Academy of Sciences*, *107*(17), 7939–7944. https://doi.org/10.1073/pnas.0904402107

Büchel, C., Price, C., & Friston, K. (1998). A multimodal language region in the ventral visual pathway. *Nature*, *394*(6690), 274–277. https://doi.org/10.1038/28389

Burton, H., Diamond, J. B., & McDermott, K. B. (2003). Dissociating Cortical Regions Activated by Semantic and Phonological Tasks: A fMRI Study in Blind and Sighted People. *Journal of Neurophysiology*, *90*(3), 1965–1982. https://doi.org/10.1152/jn.00279.2003

Cadieu, C. F., Hong, H., Yamins, D. L. K., Pinto, N., Ardila, D., Solomon, E. A., Majaj, N. J., & DiCarlo, J. J. (2014). Deep Neural Networks Rival the Representation of Primate IT Cortex for Core Visual Object Recognition. *PLoS Computational Biology*, *10*(12), e1003963. https://doi.org/10.1371/journal.pcbi.1003963

Cerpelloni, F., Collignon, O., & Op De Beeck, H. (2024). *Connecting the dots: Similar visual orthographic acquisition for Braille or line junctions*. https://doi.org/10.31234/osf.io/y3sg2

Changizi, M. A., Zhang, Q., Ye, H., & Shimojo, S. (2006). *The Structures of Letters and Symbols throughout Human History Are Selected to Match Those Found in Objects in Natural Scenes.* 23.

Cichy, R. M., Chen, Y., & Haynes, J.-D. (2011). Encoding the identity and location of objects in human LOC. *NeuroImage*, *54*(3), 2297–2307. https://doi.org/10.1016/j.neuroimage.2010.09.044

Cichy, R. M., Khosla, A., Pantazis, D., Torralba, A., & Oliva, A. (2016). Comparison of deep neural networks to spatio-temporal cortical dynamics of human visual object recognition reveals hierarchical correspondence. *Scientific Reports*, *6*(1), 27755. https://doi.org/10.1038/srep27755

Cichy, R. M., Pantazis, D., & Oliva, A. (2016). Similarity-Based Fusion of MEG and fMRI Reveals Spatio-Temporal Dynamics in Human Cortex During Visual Object Recognition. *Cerebral Cortex*, *26*(8), 3563–3579. https://doi.org/10.1093/cercor/bhw135

Cohen, L., Dehaene, S., Naccache, L., Lehéricy, S., Dehaene-Lambertz, G., Hénaff, M.-A., & Michel, F. (2000). The visual word form area: Spatial and temporal characterization of an initial stage of reading in normal subjects and posterior split-brain patients. *Brain: A Journal of Neurology*, *123*, 291–307.

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

Connolly, A. C., Guntupalli, J. S., Gors, J., Hanke, M., Halchenko, Y. O., Wu, Y.-C., Abdi, H., & Haxby, J. V. (2012). The Representation of Biological Classes in the Human Brain. *The Journal of Neuroscience*, *32*(8), 2608–2618. https://doi.org/10.1523/JNEUROSCI.5547-11.2012

Dalski, A., Kular, H., Jorgensen, J. G., Grill-Spector, K., & Grotheer, M. (2024). Both mOTS-words and pOTS-words prefer emoji stimuli over text stimuli during a lexical judgment task. *Cerebral Cortex*, *34*(8). https://doi.org/10.1093/cercor/bhae339

de Heering, A., Collignon, O., & Kolinsky, R. (2018). Blind readers break mirror invariance as sighted do. *Cortex*, *101*, 154–162. https://doi.org/10.1016/j.cortex.2018.01.002

de Heering, A., & Kolinsky, R. (2019). Braille readers break mirror invariance for both visual Braille and Latin letters. *Cognition*, *189*, 55–59. https://doi.org/10.1016/j.cognition.2019.03.012

Dębska, A., Chyl, K., Dzięgiel, G., Kacprzak, A., Łuniewska, M., Plewko, J., Marchewka, A., Grabowska, A., & Jednoróg, K. (2019). Reading and spelling skills are differentially related to phonological processing: Behavioral and fMRI study. *Developmental Cognitive Neuroscience*, *39*, 100683. https://doi.org/10.1016/j.dcn.2019.100683

Dębska, A., Wang, S., Jednoróg, K., & Pattamadilok, C. (2024). *Reading Acquisition Drives Linguistic Cross-Modal Convergence in The Left Ventral Occipitotemporal Cortex*. Neuroscience. https://doi.org/10.1101/2024.11.26.625405

Dębska, A., Wójcik, M., Chyl, K., Dzięgiel-Fivet, G., & Jednoróg, K. (2023). Beyond the Visual Word Form Area – a cognitive characterization of the left ventral occipitotemporal cortex. *Frontiers in Human Neuroscience*, *17*, 1199366. https://doi.org/10.3389/fnhum.2023.1199366

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

Dehaene, S., & Cohen, L. (2011). The unique role of the visual word form area in reading. *Trends in Cognitive Sciences*, *15*(6), 254–262. https://doi.org/10.1016/j.tics.2011.04.003

Dehaene, S., Cohen, L., Morais, J., & Kolinsky, R. (2015). Illiterate to literate: Behavioural and cerebral changes induced by reading acquisition. *Nature Reviews Neuroscience*, *16*(4), 234–244. https://doi.org/10.1038/nrn3924

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

Dehaene, S., Pegado, F., Braga, L. W., Ventura, P., Filho, G. N., Jobert, A., Dehaene-Lambertz, G., Kolinsky, R., Morais, J., & Cohen, L. (2010). How Learning to Read Changes the Cortical Networks for Vision and Language. *Science*, *330*(6009), 1359–1364. https://doi.org/10.1126/science.1194140

Dehaene-Lambertz, G., Monzalvo, K., & Dehaene, S. (2018). The emergence of the visual word form: Longitudinal evolution of category-specific ventral visual areas during reading acquisition. *PLOS Biology*, *16*(3), e2004103. https://doi.org/10.1371/journal.pbio.2004103

Downing, P. E., Jiang, Y., Shuman, M., & Kanwisher, N. (2001). A Cortical Area Selective for Visual Processing of the Human Body. *Science*, *293*(5539), 2470–2473. https://doi.org/10.1126/science.1063414

Dzięgiel-Fivet, G., Plewko, J., Szczerbiński, M., Marchewka, A., Szwed, M., & Jednoróg, K. (2021). Neural network for Braille reading and the speech-reading convergence in the blind: Similarities and differences to visual reading. *NeuroImage*, *231*, 117851. https://doi.org/10.1016/j.neuroimage.2021.117851

Epstein, R., Harris, A., Stanley, D., & Kanwisher, N. (1999). The parahippocampal place area: Recognition, navigation, or encoding?

Neuron, *23*(1), 115–125. https://doi.org/10.1016/S0896-6273(00)80758-8

Epstein, R., & Kanwisher, N. (1998). The parahippocampal place area: A cortical representation of the local visual environment. *Nature*, *392*(6676), 598–601. https://doi.org/10.1038/33402

Gaca, M., Olszewska, A. M., Droździel, D., Kulesza, A., Paplińska, M., Kossowski, B., Jednoróg, K., Matuszewski, J., Herman, A. M., & Marchewka, A. (2025). How learning to read Braille in visual and tactile domains reorganizes the sighted brain. *Frontiers in Neuroscience*, *18*, 1297344. https://doi.org/10.3389/fnins.2024.1297344

Grill-Spector, K., & Malach, R. (2004). The human visual cortex. *Annual Review of Neuroscience*, *27*(1), 649–677. https://doi.org/10.1146/annurev.neuro.27.070203.144220

Grill-Spector, K., & Weiner, K. S. (2014). The functional architecture of the ventral temporal cortex and its role in categorization. *Nature Reviews Neuroscience*, *15*(8), 536–548. https://doi.org/10.1038/nrn3747

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

Hasson, U., Levy, I., Behrmann, M., Hendler, T., & Malach, R. (2002). Eccentricity Bias as an Organizing Principle for Human High-Order Object Areas. *Neuron*, *34*(3), 479–490. https://doi.org/10.1016/S0896-6273(02)00662-1

Haupt, M., Graumann, M., Teng, S., Kaltenbach, C., & Cichy, R. M. (2024). *The transformation of sensory to perceptual braille letter representations in the visually deprived brain* [Preprint]. Neuroscience. https://doi.org/10.1101/2024.02.12.579923

Hubel, D. H., & Wiesel, T. N. (1962). Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. *The Journal of Physiology*, *160*(1), 106–154. https://doi.org/10.1113/jphysiol.1962.sp006837

Janini, D., Hamblin, C., Deza, A., & Konkle, T. (2022). General object-based features account for letter perception. *PLOS Computational Biology*, *18*(9), e1010522. https://doi.org/10.1371/journal.pcbi.1010522

Kanwisher, N., Chun, M. M., McDermott, J., & Ledden, P. J. (1996). Functional imaging of human visual recognition. *Cognitive Brain Research*, *5*(1–2), 55–67. https://doi.org/10.1016/S0926-6410(96)00041-9

Kanwisher, N., McDermott, J., & Chun, M. M. (1997). *The Fusiform Face Area: A Module in Human Extrastriate Cortex Specialized for Face Perception*. 10.

Khaligh-Razavi, S.-M., Cichy, R. M., Pantazis, D., & Oliva, A. (2018). Tracking the Spatiotemporal Neural Dynamics of Real-world Object Size and Animacy in the Human Brain. *Journal of Cognitive Neuroscience*, *30*(11), 1559–1576. https://doi.org/10.1162/jocn_a_01290

Kim, J. G., Biederman, I., Lescroart, M. D., & Hayworth, K. J. (2009). Adaptation to objects in the lateral occipital complex (LOC): Shape or semantics? *Vision Research*, *49*(18), 2297–2305. https://doi.org/10.1016/j.visres.2009.06.020

Kolinsky, R., Verhaeghe, A., Fernandes, T., Mengarda, E. J., Grimm-Cabral, L., & Morais, J. (2011). Enantiomorphy through the looking glass: Literacy effects on mirror-image discrimination. *Journal of Experimental Psychology: General*, *140*(2), 210–238. https://doi.org/10.1037/a0022168

Korczyk, M., Rączy, K., & Szwed, M. (2024). Mirror-invariance is not exclusively visual but extends to touch. *Scientific Reports*, *14*(1), 31094. https://doi.org/10.1038/s41598-024-82350-6

Kourtzi, Z. (2003). Representation of the Perceived 3-D Object Shape in the Human Lateral Occipital Complex. *Cerebral Cortex*, *13*(9), 911–920. https://doi.org/10.1093/cercor/13.9.911

Kourtzi, Z., & Kanwisher, N. (2000). Cortical Regions Involved in Perceiving Object Shape. *The Journal of Neuroscience*, *20*(9), 3310–3318. https://doi.org/10.1523/JNEUROSCI.20-09-03310.2000

Kourtzi, Z., & Kanwisher, N. (2001). Representation of Perceived Object Shape by the Human Lateral Occipital Complex. *Science*, *293*(5534), 1506–1509. https://doi.org/10.1126/science.1061133

Kriegeskorte, N., Mur, M., & Bandettini, P. (2008). Representational similarity analysis—Connecting the branches of systems neuroscience. *Frontiers in Systems Neuroscience*, *2*(NOV), 1–28. https://doi.org/10.3389/neuro.06.004.2008

Kriegeskorte, N., Mur, M., Ruff, D. A., Kiani, R., Bodurka, J., Esteky, H., Tanaka, K., & Bandettini, P. A. (2008). Matching Categorical Object Representations in Inferior Temporal Cortex of Man and Monkey. *Neuron*, *60*(6), 1126–1141. https://doi.org/10.1016/j.neuron.2008.10.043

Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2017). ImageNet classification with deep convolutional neural networks. *Communications of the ACM*, *60*(6), 84–90. https://doi.org/10.1145/3065386

Kubilius, J., Schrimpf, M., Nayebi, A., Bear, D., Yamins, D. L. K., & DiCarlo, J. J. (2018). *CORnet: Modeling the Neural Mechanisms of Core Object Recognition*. Neuroscience. https://doi.org/10.1101/408385

Lerma-Usabiaga, G., Carreiras, M., & Paz-Alonso, P. M. (2018). Converging evidence for functional and structural segregation within the left ventral occipitotemporal cortex in reading. *Proceedings of the National Academy of Sciences*, *115*(42). https://doi.org/10.1073/pnas.1803003115

Li, J., Hiersche, K. J., & Saygin, Z. M. (2024). Demystifying the Visual Word Form Area: Visual and Nonvisual Response Properties of Ventral Temporal Cortex with precision fMRI. *bioRxiv*. https://doi.org/10.1101/2023.06.15.544824

Li, J., Osher, D. E., Hansen, H. A., & Saygin, Z. M. (2020). Innate connectivity patterns drive the development of the visual word form area. *Scientific Reports*, *10*(1), 18039. https://doi.org/10.1038/s41598-020-75015-7

Liu, J., Li, J., Feng, L., Li, L., Tian, J., & Lee, K. (2014). Seeing Jesus in toast: Neural and behavioral correlates of face pareidolia. *Cortex*, *53*, 60–77. https://doi.org/10.1016/j.cortex.2014.01.013

Logothetis, N. K., & Sheinberg, D. L. (1996). *Visual Object Recognition*.

Malach, R., Reppas, J. B., Benson, R. R., Kwong, K. K., Jiang, H., Kennedy, W. A., Ledden, P. J., Brady, T. J., Rosen, B. R., & Tootell, R. B. (1995). Object-related activity revealed by functional magnetic resonance imaging in human occipital cortex. *Proceedings of the National Academy of*

Sciences, *92*(18), 8135–8139. https://doi.org/10.1073/pnas.92.18.8135

Martin, L., Durisko, C., Moore, M. W., Coutanche, M. N., Chen, D., & Fiez, J. A. (2019). The VWFA Is the Home of Orthographic Learning When Houses Are Used as Letters. *Eneuro*, *6*(1), ENEURO.0425-17.2019. https://doi.org/10.1523/ENEURO.0425-17.2019

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

Moore, M. W., Durisko, C., Perfetti, C. A., & Fiez, J. A. (2014). Learning to Read an Alphabet of Human Faces Produces Left-lateralized Training Effects in the Fusiform Gyrus. *Journal of Cognitive Neuroscience*, *26*(4), 896–913. https://doi.org/10.1162/jocn_a_00506

Op de Beeck, H. P., DiCarlo, J. J., Goense, J. B. M., Grill-Spector, K., Papanastassiou, A., Tanifuji, M., & Tsao, D. Y. (2008). Fine-Scale Spatial Organization of Face and Object Selectivity in the Temporal Lobe: Do Functional Magnetic Resonance Imaging, Optical Imaging, and Electrophysiology Agree? *Journal of Neuroscience*, *28*(46), 11796–11801. https://doi.org/10.1523/JNEUROSCI.3799-08.2008

Op de Beeck, H. P., Pillet, I., & Ritchie, J. B. (2019). Factors Determining Where Category-Selective Areas Emerge in Visual Cortex. *Trends in Cognitive Sciences*, *23*(9), 784–797. https://doi.org/10.1016/j.tics.2019.06.006

Pattamadilok, C., Planton, S., & Bonnard, M. (2019). Spoken language coding neurons in the Visual Word Form Area: Evidence from a TMS

adaptation paradigm. *NeuroImage*, *186*, 278–285. https://doi.org/10.1016/j.neuroimage.2018.11.014

Pegado, F., Nakamura, K., Cohen, L., & Dehaene, S. (2011). Breaking the symmetry: Mirror discrimination for single letters but not for pictures in the Visual Word Form Area. *NeuroImage*, *55*(2), 742–749. https://doi.org/10.1016/j.neuroimage.2010.11.043

Pegado, F., Nakamura, K., & Hannagan, T. (2014). How does literacy break mirror invariance in the visual system? *Frontiers in Psychology*, *5*. https://doi.org/10.3389/fpsyg.2014.00703

Pillet, I., Cerrahoğlu, B., Philips, R. V., Dumoulin, S., & De Beeck, H. O. (2024). The position of visual word forms in the anatomical and representational space of visual categories in occipitotemporal cortex. *Imaging Neuroscience*, *2*, 1–28. https://doi.org/10.1162/imag_a_00196

Pillet, I., Cerrahoğlu, B., Philips, R. V., Dumoulin, S., & Op De Beeck, H. (2024). A 7T fMRI investigation of hand and tool areas in the lateral and ventral occipitotemporal cortex. *PLOS ONE*, *19*(11), e0308565. https://doi.org/10.1371/journal.pone.0308565

Planton, S., Chanoine, V., Sein, J., Anton, J.-L., Nazarian, B., Pallier, C., & Pattamadilok, C. (2019). Top-down activation of the visuo-orthographic system during spoken sentence processing. *NeuroImage*, *202*, 116135. https://doi.org/10.1016/j.neuroimage.2019.116135

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

Price, C. J., & Devlin, J. T. (2011). The Interactive Account of ventral occipitotemporal contributions to reading. *Trends in Cognitive Sciences*, *15*(6), 246–253. https://doi.org/10.1016/j.tics.2011.04.001

Ptito, M., Schneider, F. C. G., Paulson, O. B., & Kupers, R. (2008). Alterations of the visual pathways in congenital blindness. *Experimental Brain Research*, *187*(1), 41–49. https://doi.org/10.1007/s00221-008-1273-4

Rajalingham, R., Kar, K., Sanghavi, S., Dehaene, S., & DiCarlo, J. J. (2020). The inferior temporal cortex is a potential cortical precursor of orthographic processing in untrained monkeys. *Nature Communications*, *11*(1), 3886. https://doi.org/10.1038/s41467-020-17714-3

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

Riesenhuber, M., & Poggio, T. (1999). Hierarchical models of object recognition in cortex. *Nature Neuroscience*, *2*(11), 1019–1025. https://doi.org/10.1038/14819

Riesenhuber, M., & Poggio, T. (2002). Neural mechanisms of object recognition. *Current Opinion in Neurobiology*, *12*(2), 162–168. https://doi.org/10.1016/S0959-4388(02)00304-5

Roberts, D. J., Woollams, A. M., Kim, E., Beeson, P. M., Rapcsak, S. Z., & Lambon Ralph, M. A. (2013). Efficient Visual Object and Word Recognition Relies on High Spatial Frequency Coding in the Left Posterior Fusiform Gyrus: Evidence from a Case-Series of Patients with Ventral Occipito-Temporal Cortex Damage. *Cerebral Cortex*, *23*(11), 2568–2580. https://doi.org/10.1093/cercor/bhs224

Russakovsky, O., Deng, J., Su, H., Krause, J., Satheesh, S., Ma, S., Huang, Z., Karpathy, A., Khosla, A., Bernstein, M., Berg, A. C., & Fei-Fei, L. (2015). ImageNet Large Scale Visual Recognition Challenge. *International Journal of Computer Vision*, *115*(3), 211–252. https://doi.org/10.1007/s11263-015-0816-y

Sadato, N., Okada, T., Kubota, K., & Yonekura, Y. (2004). Tactile discrimination activates the visual cortex of the recently blind naive to Braille: A functional magnetic resonance imaging study in humans. *Neuroscience Letters*, *359*(1–2), 49–52. https://doi.org/10.1016/j.neulet.2004.02.005

Sadato, N., Pascual-Leone, A., Grafman, J., Deiber, M.-P., Ibanez, V., & Hallett, M. (1998). *Neural networks for Braille reading by the blind*.

Saygin, Z. M., Osher, D. E., Norton, E. S., Youssoufian, D. A., Beach, S. D., Feather, J., Gaab, N., Gabrieli, J. D. E., & Kanwisher, N. (2016). Connectivity precedes function in the development of the visual word form area. *Nature Neuroscience*, *19*(9), 1250–1255. https://doi.org/10.1038/nn.4354

Seeliger, K., Fritsche, M., Güçlü, U., Schoenmakers, S., Schoffelen, J.-M., Bosch, S. E., & Van Gerven, M. A. J. (2018). Convolutional neural network-based encoding and decoding of visual object recognition in space and time. *NeuroImage*, *180*, 253–266. https://doi.org/10.1016/j.neuroimage.2017.07.018

Siuda-Krzywicka, K., Bola, Ł., Paplińska, M., Sumera, E., Jednoróg, K., Marchewka, A., Śliwińska, M. W., Amedi, A., & Szwed, M. (2016). Massive cortical reorganization in sighted Braille readers. *eLife*, *5*, e10762. https://doi.org/10.7554/eLife.10762

Szwed, M., Cohen, L., Qiao, E., & Dehaene, S. (2009). The role of invariant line junctions in object and visual word recognition. *Vision Research*, *49*(7), 718–725. https://doi.org/10.1016/j.visres.2009.01.003

Szwed, M., Dehaene, S., Kleinschmidt, A., Eger, E., Valabrègue, R., Amadon, A., & Cohen, L. (2011). Specialization for written words over objects in the visual cortex. *NeuroImage*, *56*(1), 330–344. https://doi.org/10.1016/j.neuroimage.2011.01.073

Testolin, A., Stoianov, I., & Zorzi, M. (2017). Letter perception emerges from unsupervised deep learning and recycling of natural image features. *Nature Human Behaviour*, *1*(9), 657–664. https://doi.org/10.1038/s41562-017-0186-2

Van Atteveldt, N., Formisano, E., Goebel, R., & Blomert, L. (2004). Integration of Letters and Speech Sounds in the Human Brain. *Neuron*, *43*(2), 271–282. https://doi.org/10.1016/j.neuron.2004.06.025

Van Audenhaege, A., Mattioni, S., Cerpelloni, F., Remi, G., Arnaud, S., & Collignon, O. (2024). *Phonological representations of auditory and visual speech in the occipito-temporal cortex and beyond*. https://doi.org/10.1101/2024.07.25.605084

Vinckier, F., Dehaene, S., Jobert, A., Dubus, J. P., Sigman, M., & Cohen, L. (2007). Hierarchical Coding of Letter Strings in the Ventral Stream: Dissecting the Inner Organization of the Visual Word-Form System. *Neuron*, *55*(1), 143–156. https://doi.org/10.1016/j.neuron.2007.05.031

Vogel, A. C., Petersen, S. E., & Schlaggar, B. L. (2012). The Left Occipitotemporal Cortex Does Not Show Preferential Activity for Words. *Cerebral Cortex*, *22*(12), 2715–2732. https://doi.org/10.1093/cercor/bhr295

Wang, S., Planton, S., Chanoine, V., Sein, J., Anton, J.-L., Nazarian, B., Dubarry, A.-S., Pallier, C., & Pattamadilok, C. (2022). Graph theoretical analysis reveals the functional role of the left ventral occipito-temporal cortex in speech processing. *Scientific Reports*, *12*(1), 20028. https://doi.org/10.1038/s41598-022-24056-1

Wang, X., Xu, Y., Wang, Y., Zeng, Y., Zhang, J., Ling, Z., & Bi, Y. (2018). Representational similarity analysis reveals task-dependent semantic influence of the visual word form area. *Scientific Reports*, *8*(1), 3047. https://doi.org/10.1038/s41598-018-21062-0

White, A. L., Kay, K. N., Tang, K. A., & Yeatman, J. D. (2023). Engaging in word recognition elicits highly specific modulations in visual cortex. *Current Biology*, *33*(7), 1308-1320.e5. https://doi.org/10.1016/j.cub.2023.02.042

Xue, G., & Poldrack, R. A. (2007). The Neural Substrates of Visual Perceptual Learning of Words: Implications for the Visual Word Form Area Hypothesis. *Journal of Cognitive Neuroscience*, *19*(10), 1643–1655. https://doi.org/10.1162/jocn.2007.19.10.1643

Yamins, D. L. K., Hong, H., Cadieu, C. F., Solomon, E. A., Seibert, D., & DiCarlo, J. J. (2014). Performance-optimized hierarchical models predict neural responses in higher visual cortex. *Proceedings of the National Academy of Sciences*, *111*(23), 8619–8624. https://doi.org/10.1073/pnas.1403112111

Yargholi, E., & Op De Beeck, H. (2023). Category Trumps Shape as an Organizational Principle of Object Space in the Human Occipitotemporal Cortex. *The Journal of Neuroscience*, *43*(16), 2960–2972. https://doi.org/10.1523/JNEUROSCI.2179-22.2023

Yeatman, J. D., Rauschecker, A. M., & Wandell, B. A. (2013). Anatomy of the visual word form area: Adjacent cortical circuits and long-range white matter connections. *Brain and Language*, *125*(2), 146–155. https://doi.org/10.1016/j.bandl.2012.04.010

Yeatman, J. D., & White, A. L. (2021). Reading: The Confluence of Vision and Language. *Annual Review of Vision Science*, *7*(1), 487–517. https://doi.org/10.1146/annurev-vision-093019-113509

Zhan, M., Pallier, C., Agrawal, A., Dehaene, S., & Cohen, L. (2023). Does the visual word form area split in bilingual readers? A millimeter-scale 7-T fMRI study. *Science Advances*.

Chapter 2. Similar visual orthographic acquisition for Braille or line junctions

During the PhD, Filippo Cerpelloni was the daily supervisor of two master theses (Emma Gastoud, UCLouvain; Tom Poel, KU Leuven) that collaborated to behavioural testing of individuals learning either Braille or Line Braille (Chapter 2). Emma Gastoud explored a pilot study, in a different language and with a restricted number of subjects (https://hdl.handle.net/2078.2/29290). Tom Poel was involved in the final experiment, helping in the development of pilot experiments and in the initial phase of the main study. These theses explored some initial and partial aspects, further developed and strongly extended by Filippo Cerpelloni.

This chapter as provided in the thesis is published as a pre-print on PsyArXiv and is currently under review in the journal "Psychonomic Bulletin and Reviews". Reference: Cerpelloni, F., Collignon, O.*, & Op de Beeck, H.* (2024, November 19). Connecting the dots: similar visual orthographic acquisition for Braille or line junctions. https://doi.org/10.31234/osf.io/y3sg2_v1

Abstract

Most written systems share basic shape features with natural objects such as line junctions, a commonality thought to be at the basis of fluent reading. Studies that compared reading acquisition for different scripts used line-based scripts and investigated them in terms of complexity, novelty, or associations of stimuli from different categories. Here, we directly compared a novel, custom-made, line-based script derived from Braille (Line Braille) to visual Braille, a script that only includes patterns of dots and no line junctions. For four consecutive days, two groups of participants underwent online training, during which they first mapped new letters onto the Latin alphabet and then trained on full words. Each day, participants were tested by asking them to transcribe a set of stimuli (words and pseudowords) from the novel to the Latin script. Across sessions, we found no significant differences between scripts in the overall transcription accuracy nor in the time required to transcribe words. We only found a small delay in the learning trajectory of the Braille group represented through an interaction between group and session in overall accuracy, and slightly higher sensitivity for stimulus length in the Braille group. Overall, these results show that line junctions only provide a small and temporary benefit when learning a new script, contrasting with the idea that line junctions is a core visual features for learning orthography.

Introduction

Visual object recognition relies on the integration of visual inputs into progressively more complex objects (Logothetis & Sheinberg, 1996; Riesenhuber & Poggio, 1999, 2002). One particular case of object recognition concerns reading. Behaviourally, reading consist in the recognition of letters and words, attributing linguistic value to an arbitrary visual code. The main theories behind reading acquisition in the brain pose that regions in the ventral visual stream are co-opted to process written stimuli (Cohen et al., 2000, 2002), integrating the features that compose letters into progressively complex stimuli, culminating into words (Dehaene et al., 2005; Dehaene-Lambertz et al., 2018; Vinckier et al., 2007). Literature on this topic focused on the importance and essentiality of the lines and the junctions composing such stimuli, following studies that showed how most written scripts share the same basic properties, line junctions, and that those properties are derived from natural objects (Changizi et al., 2006), thus hypothesising an evolution of scripts to accommodate the existing preference of the brain for the integration of line junctions.

Prominent studies showed for example that, similarly to line drawings, disrupting the line vertices (junctions) of letters results in higher reaction times and more naming errors (Szwed et al., 2009) and that this behaviour is reflected in brain activations (Szwed et al., 2011). In contrast, several behavioural and neuroimaging studies support an alternative account of reading, focused on the connections between the visual system and language areas in the brain (Saygin et al., 2016). In particular, several scripts have been associated to VWFA activation (houses, Martin et al. (2019); faces, Moore et

al. (2014); Hebrew or Korean, Baker et al. (2007); Xue & Poldrack (2007)). Putting aside the specifics of activations in VWFA, all these studies feature stimuli that, for how complex and novel they can be, are made out of lines and junctions, visual units that are common to most scripts (Changizi et al., 2006). Under this consideration, all these studies can be seen as confirmations of the adaptability of the visual reading system to complex line-based stimuli.

Little evidence is present, instead, for what happens when the line junctions are missing. A unique case is presented by visual reading of Braille. In Braille, letters consist of a combination of 1 to 6 dots. With its uniquely different structure, it poses an interesting paradigm to study visual learning and reading acquisition. While Braille is not made of the canonical line junctions (Changizi et al., 2006), in the case of western Latin-based scripts, it maps into the same linguistic properties (orthography, phonology) of the script it mirrors. Previous work has already investigated tactile Braille reading in blind (Büchel et al., 1998; Reich et al., 2011) and sighted individuals (Gaca et al., 2024; Siuda-Krzywicka et al., 2016), and visual Braille reading in both behavioural (Bola et al., 2017) and neuroimaging studies (Cerpelloni et al., 2024). A direct comparison of visual Braille and line-based scripts is however missing. Among these studies, only Bola and colleagues (Bola et al., 2017) compared a training on visual Braille to a training on another script, Cyrillic, leading to the conclusion that line junctions may be needed for fluent reading. Their study shows how a line-based script leads to higher accuracy and lower reaction times in a lexical decision task compared to visual Braille. This study, however, compares two groups who underwent different types of training: those who learnt Cyrillic attended a Russian language course, while those who learnt Braille were

assigned with individual work to be performed in tactile Braille and checked visually. Not only the trainings were different, but visual Braille reading was not trained directly and its effect are not distinguishable from the effect of tactile Braille reading. Moreover, while Cyrillic is visually dissimilar from the Polish (the language of the participants of the study), and while participants were non-experts in Russian (the Cyrillic language learnt), they might have been previously exposed to it. Cyrillic characters partially overlap with Latin ones (letters A, B, I, P, C, H, X, T, are all occurring in both scripts), in some cases with similar grapheme-phoneme mapping (e.g. A, O, T), and such overlap may have posed an advantage for Cyrillic over Braille script.

To overcome these limitations, in our study we compare behavioural responses for the same training paradigm applied to two highly similar scripts: visual Braille and a custom script where Braille dots are connected to form line junctions. By reducing the differences between the training paradigms and between the scripts that are compared, we can more effectively identify whether the hypothesized advantage of lines is indeed present and, if so, to what extent.

We compared the behavioural responses of two groups of participants learning either script in the same online training setting. We observed that overall accuracy and time to provide the answer during a transcription of words from the novel to the native alphabet does not depend on which script is learnt, meaning that the visual features of the script (the only difference between them) does not pose an advantage. The only consistent but temporary advantage we observed was a faster learning in the line-based script, reflected in a significant interaction between script and session. In

general, we show a limited and temporary advantage to learn a line-based script to one that does not present such features. These findings illustrate the flexibility of the visual system to adapt to peculiar stimuli that lack the typical features of objects and scripts.

Methods

Participants

A total of eighty participants (53 Females, mean age = 22.5 ± 3.97) completed the experiment within the time windows requested a priori (one session per day) and were included in the analyses. An additional twenty-one participants were recruited but excluded prior to the start of analyses due to different reasons (fourteen did not complete the experiment, five did not start, and two did not complete the experiment in the required time window). Participants were assigned a script to learn at the beginning of the experiment, resulting in two matched groups which learnt either Braille (N = 40, 26 Females, mean age = 22.52 ± 3.65) or Line Braille (N = 40, 27 Females, mean age = 22.48 ± 4.29). The protocol was approved by the ethics committee at KU Leuven and all participants gave explicit informed consent prior to the start of the first training session.

Stimuli

To compose the new alphabets, we relied on the Braille script and joined the dots forming each letter into continuous lines. For this reason, we adapted the Braille script and modified some characters. In particular, we changed the

character for the letter ‘a’ from ‘⠁’ to ‘⠤’ to allow for a line counterpart in in the new line alphabet; and we changed the character for the letter ‘k’ from ‘⠅’ to ‘⠢’ to avoid an overlap with the line letter ‘l’ (‘⠇’). We used letter stimuli created for a previous pilot investigation (Gastoud, 2022) in the French script. We created the Line Braille condition by joining the dots to create the shortest path possible and to avoid similarity with Latin script letters, where possible. To select the words for the training set, presented in sessions two, three, and four, we relied on the Dutch Lexicon Project 2 (Brysbaert et al., 2016). All the words we selected had high frequency (> 2/million) and word prevalence (higher than 1.65), lengths between four and eight characters and between two and six phonemes. The final set consisted of two hundred words, equally distributed based on the length of the words. From this set, we selected twenty different words for each of the three words training session (sixty words in total) to perform an interim assessment of the transcription accuracy during training. We modelled the stimulus set of the test words based on the work of Martin and colleagues (Martin et al., 2019), where the authors presented novel words, previously seen words, and pseudowords. To choose the novel words, we adopted the same criteria used to select the training words. We chose the seen words from the training set, randomly selecting words that were not part of the interim assessment. Lastly, we choose the pseudowords from Wuggy (Keuleers & Brysbaert, 2010), a multilingual pseudoword generator. For each subset, we picked eighty stimuli of matched length, from four to eight characters. We then assigned twenty stimuli to each testing session, balancing the lengths of the stimuli across sessions.

Experimental design

We presented the experiments to participants using PsychoPy (Peirce, 2007, 2008) through Pavlovia.org, a web-based platform for the presentation of experiments. With the exceptions of the script learnt and the random presentation order of the stimuli, all the participants underwent the same training procedure (Figure 1B). All included participants performed the experiment sessions on four consecutive days. Although the experiment was self-paced, and resulted in different durations between participants, we did not find significant differences across scripts (Figure 1C) in the duration of each session (session 1: $t_{(78)} = 0.02$, $p_{Bonf} = 1$; session 2: $t_{(78)} = 1.16$, $p_{Bonf} = 0.99$; session 3: $t_{(78)} = 1.25$, $p_{Bonf} = 0.86$; session 4: $t_{(78)} = 1.38$, $p_{Bonf} = 0.69$).

On the first session, participants had to read and accept the informed consent form to continue and start the experiment. They were then prompted to learn the new alphabet letters (Figure 1D). Each letter was presented at the center of the screen, surrounded by a frame to highlight the relative position of the features. After 1.5 seconds, the corresponding letter in the Latin script appeared. Participants could study the correspondence between scripts without a time limit, but for a minimum time of 1.5 seconds and could progress in the learning by a mouse press. The whole alphabet of new stimuli (twenty-six letters) was presented for seven repetitions, with randomized letter order within one repetition. The training of letters lasted for 21 (± 7) minutes on average. Participants were then prompted to take a pause of approximately five minutes before starting the test, and were provided with instructions on how to perform the test.

The test phase's methodology was identical, except for the stimuli, in each session (Figure 1F). Participants were presented with a string of characters (a word or a pseudoword) for 6 seconds. After that time elapsed, the stimulus

disappeared and participants had to provide a written transliteration of the stimulus in the Latin script. To submit the answer, participants had to perform a mouse press. This procedure allowed us to register not only the accuracy of the transliteration, but also the time required to provide the answer, adding a measure of learning speed, or efficiency. After an answer was submitted, participants moved to the next test item.

On the second session, participants underwent first one repetition of the letters training, to review the correspondence. This phase lasted for 2.3 (± 0.8) minutes. After the review of letters, instructions for the training experiment were provided and participants could start the session. Each item of the training set was a full word presented at the center of the screen. In most trials, after 2.5 seconds participants were able to press a key to show the corresponding word in the Latin script, and only after the answer was revealed they could continue to the next trial. In the case of one item selected for the interim assessment (10% of the trials, 20 words), participants had to input a string of text to reveal the answer and to be allowed to continue the experiment (Figure 1E). We opted to add these trials for two reasons: to obtain a measure of the pace of learning, and to ensure attentive participation in the experiment. After the training phase, which lasted on average 32.95 (± 14.36) minutes, participants were again prompted to take a break and perform another test session, in the same manner as the previous day. On the third and fourth sessions, participants performed again the words training, under the procedure but with different trials assessed, and underwent the same testing phases, with new sets of stimuli.

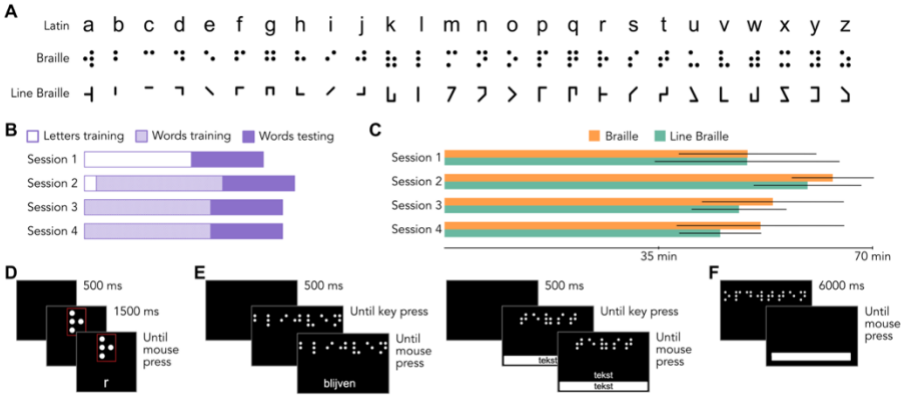

Figure 2 – experimental design. (A) Letters used in the experiment: Latin, Braille, Line Braille. (B) Schematic view of the training experiment. In the first session, participants saw seven cycles of letters in the new alphabet with the relative transcription and then moved to a test phase on the transcription of full words and pseudo-words. In all the subsequent session, training was on full words, while testing was on words and pseudo-words. At the start of the second session, participants saw another cycle of letters before starting to learn full words. (C) Average time spent in each session, divided by group. Statistical tests do not show significant differences between groups in terms of time spent on each session. (D) Trial of letters training. Each trial started with a 500ms black screen, after which a letter in the novel alphabet was presented inside a red outline. The outline was added to provide information about the vertical position of the dots / lines. The novel letter remained on screen until participants pressed a key on their keyboard to show the corresponding Latin letter. After the transcription was presented, participants could advance via a mouse click. (E) Trials of words training. In the majority of trials (90%, left), participants were required to read the novel word presented on screen after a 500ms interval. Only after 3500ms, they could show the transcription and move forward. In 10% of the trials (right), a written answer was required. Participants had to input the transcription and could check their answer by a key press. Only after the key press, they could advance to the next item. (F) Test trial. Each test trial started with a six seconds presentation of the test stimulus, after which the stimulus would disappear and an answer box would appear for the participant to input the correct transcription. The participant could then move to the next item by clicking with the mouse to submit the answer. (E)

Statistical analyses

To ensure that the training paradigms are comparable, we performed four independent sample t-tests on the durations of the sessions across scripts, and corrected for multiple comparisons using Bonferroni correction. To test whether one script was easier to learn, we extracted several measures of

accuracy and timing from the results of the three sessions on words training and the four test sessions. From the test results, we computed the accuracy for each participant at each session. We also computed the writing time, the time required to provide the written answer. We measured such time by the time between the first keypress and the last keypress on the keyboard. In this computation, we considered as an outlier any case where the total time was higher than 60 seconds, or when the time between two keypresses was higher than 30 seconds. We chose this measure to exclude potential outlier cases where the participant, in order to pause the experiment, would leave the completed answer on screen before moving to the next item. We additionally computed accuracy during the training interim assessment and the time required to process the words not assessed. In this case, we considered as outlier any trial that required more than 60 seconds, for the same reason described above.

For each behavioural measure, we performed repeated measures ANOVA on the test (2 scripts * 4 test sessions) and on the training (2 scripts * 3 training sessions) results. When an interaction effect was found, we performed post-hoc tests by means of four t-tests on independent samples on the differences between scripts at each session. We corrected those test for multiple comparisons using Bonferroni correction. To further interpret whether the lack of significant effect corresponds to a lack of finding or to the finding of a lack of effect, we computed Bayes Factors (Kass & Raftery, 1995). To investigate the role played by the linguistic properties of the stimuli learnt, we computed correlations between the results from each participant and the stimuli's length, frequency, and orthographic neighbours. To identify significant correlations between behaviour and linguistic properties, we performed, for each

behaviour-linguistic correlation, a one sample t-test against chance (zero) on the average correlation per participant. We then performed repeated measures ANOVAs on the individual correlations to examine differences between scripts and sessions. For all the repeated measures ANOVAs, we estimated effect size through partial eta-squared (η^2_p).

Results

Transcription accuracy during testing

To assess whether the two scripts can be transcribed from the novel to the Latin alphabet in a similar manner, the first and most straightforward measure we obtained was transcription accuracy (Figure 2A). Overall, across the four testing sessions, we observed a main effect of session ($F_{(3, 234)} = 236.231$, $p < 0.001$, $\eta^2_p = 0.75$), indicating that the participants in both groups improve over time. We did not observe a main effect of the script tested ($F_{(1, 78)} = 2.312$, $p = 0.13$, $\eta^2_p = 0.03$). This lack of significance shows that, as early as after a few hours of training divided in four consecutive days, the two scripts reach comparable levels of accuracy and are therefore assimilated in a similar manner. Interestingly, we do observe a significant interaction between sessions and scripts ($F_{(3, 234)} = 3.678$, $p = 0.012$, $\eta^2_p = 0.05$), which highlights a difference between the learning curves. To clarify the nature of such interaction, we performed post-hoc analyses on each session to identify where a difference between script could be present. We observed a difference only in the first session ($t_{(78)} = -2.39$, $p_{uncorr} = 0.018$) and not in the following ones (sessions 2, 3, 4: all $p_{uncorr} > 0.26$), supporting the interaction and the lack of

main differences between scripts. However, no significant test survives Bonferroni correction (session 1: p_{Bonf} = 0.076; sessions 2, 3, 4: all p_{Bonf} = 1). Such limited interaction is supported by Bayes Factors, which show an initial limited evidence (session 1: BF = 2.69, measurement error err. < 0.001) for the difference between scripts, but increasing evidence for the null hypothesis in the following sessions (session 2: BF = 0.41, err. = 0.0015; session 3: BF = 0.35, err. = 0.0016; session 4: BF = 0.26, err. = 0.0017).

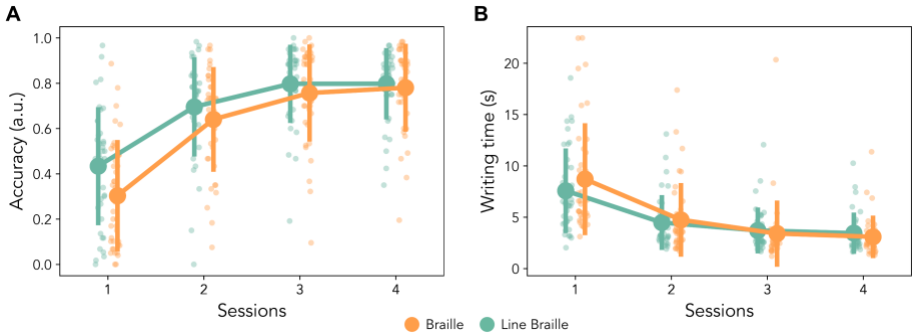

Figure 3 – Accuracy and time to write the transliteration during test. (A) Accuracy in the transcriptions. A main effect of session indicates that the transcriptions improve over time, while the lack of group differences indicates that the two scripts are learnt in a similar, comparable manner. There was a significant interaction between session and script. (B) Writing time to provide the transcription. The speed of transcription improves during the training, without differences between the scripts, nor interactions between scripts and sessions. Error bars represent standard error.

Writing time during word testing

Similarly to accuracy, we investigated the time required to transcribe the (pseudo)words presented into their native Latin script (Figure 2B). We performed a repeated measures ANOVA (rmANOVA) on sessions * script and found only a main effect of session ($F_{(3, 234)}$ = 98.451, p < 0.001, η^2_p = 0.56), indicating a progressive improvement and efficiency in the visual learning task.

We did not find significant differences between script ($F_{(1, 78)} = 0.099$, $p = 0.75$, $\eta^2_p = 0.001$) nor interaction effects between the variables ($F_{(3, 234)} = 2.334$, $p = 0.075$, $\eta^2_p = 0.03$). Overall, the learning progression seems to be equal between the two scripts as far as it can be assessed through the time that participants take to write the words in full. Although not significant, we explored the interaction trend between scripts and sessions. Post-hoc t-tests on the script difference reveal no significant differences at any point (all sessions $p_{uncorr} > 0.28$, $p_{Bonf} = 1$), with Bayes Factors pointing in favour of the null hypothesis (session 1: BF = 0.39, err. < 0.001; session 2: BF = 0.24, err. < 0.001; session 3: BF = 0.26, err. < 0.001; session 4: BF = 0.32, err. < 0.001).

Influences of language statistics

To assess whether the differences observed were influenced by statistics of the stimuli presented, we performed correlations between the accuracy measures and writing time with the linguistic statistics of the word and pseudowords used in the test sets (Figure 3). To extract language statistics of the novel and seen words subsets, we relied on the Dutch Lexicon Project (Brysbaert et al., 2016) to extract information about the length, the frequency, and the orthographic neighbours (old20) of a given word. For the pseudo-words, we could only compute a measure of the length of the word, therefore the correlations with length will refer to the whole corpus of test stimuli presented (N = 60 per session), while correlations with frequency and orthographic neighbours will be limited to the subset of real words (N = 40 per session).

We observe, in all the cases, overall significant correlations between the behavioural measurements and the linguistic statistics of the stimuli. Accuracy

correlates negatively with the length of the stimulus, with less accurate responses for longer stimuli ($r = -0.23$, $t_{(79)} = -21.30$, $p_{Bonf} < 0.001$) and with orthographic neighbours of the target words ($r = -0.21$, $t_{(78)} = -19.42$, $p_{Bonf} < 0.001$), meaning that words with more orthographic neighbours lead to more errors, but positively with word frequency ($r = 0.08$, $t_{(78)} = 12.02$, $p_{Bonf} < 0.001$), which indicates that more frequent words are easier to transcribe than non-frequent ones. As for the time to write the answers, we observe opposite trends (timing-length: $r = 0.30$, $t_{(79)} = 25.58$, $p_{Bonf} < 0.001$; timing-ort. neighbours: $r = 0.32$, $t_{(79)} = 25.69$, $p_{Bonf} < 0.001$; timing-frequency: $r = -0.12$, $t_{(79)} = -17.42$, $p_{Bonf} < 0.001$) but with similar interpretations: longer stimuli induce longer writing time, and words with more similar orthographic neighbours lead to more time to complete the word. Lastly, more frequent words required less time to be written.

More in depth, we compared the influence of each linguistic statistic over the progression of the two groups across test sessions. Regarding the correlation between transcription accuracy and stimulus length, we found a significant effect of session ($F_{(3, 216)} = 8.52$, $p < 0.001$, $\eta^2_p = 0.11$) and a difference between the scripts across the whole training ($F_{(1, 72)} = 6.06$, $p = 0.016$, $\eta^2_p = 0.078$). Such difference and its effect size indicate that participants who learnt Braille were slightly more sensitive to the length of the word or pseudoword presented (i.e. more prone to errors if the word was longer) than the group who learnt Line Braille. No interaction is present between script and session ($F_{(3, 216)} = 0.22$, $p = 0.88$, $\eta^2_p = 0.003$).

In the correlations between accuracy and orthographic neighbours (session: $F_{(3, 189)} = 14.13$, $p < 0.001$, $\eta^2_p = 0.97$; script: $F_{(1, 63)} = 0.65$, $p = 0.42$, $\eta^2_p = 0.01$; script*session: $F_{(3, 189)} = 1.00$, $p = 0.39$, $\eta^2_p = 0.02$), and between accuracy and

word frequency (session: $F_{(3, 189)} = 4.21$, $p = 0.006$, $\eta^2_p = 0.06$; script: $F_{(1, 63)} = 0.04$, $p = 0.83$, $\eta^2_p < 0.001$; script*session: $F_{(3, 189)} = 0.11$, $p = 0.95$, $\eta^2_p = 0.001$) we only observe main effects of session and no script differences, which indicate similar progression in learning across days of training. Between writing time and linguistic statistics (Figure 3B), we found significant main effects of the training sessions and no difference between scripts. This is valid for the correlations between writing time and stimulus length (session: $F_{(3, 234)} = 6.78$, $p < 0.001$, $\eta^2_p = 0.08$; script: $F_{(1, 78)} = 3.47$, $p = 0.07$, $\eta^2_p = 0.04$; script*session: $F_{(3, 234)} = 1.86$, $p = 0.13$, $\eta^2_p = 0.02$), word orthographic neighbours (session: $F_{(3, 234)} = 7.13$, $p < 0.001$, $\eta^2_p = 0.08$; script: $F_{(1, 78)} = 0.85$, $p = 0.36$, $\eta^2_p = 0.01$; script*session: $F_{(3, 234)} = 0.73$, $p = 0.53$, $\eta^2_p = 0.01$), and word frequency (session: $F_{(3, 234)} = 10.40$, $p < 0.001$, $\eta^2_p = 0.11$; script: $F_{(1, 78)} = 3.86$, $p = 0.053$, $\eta^2_p = 0.05$; script*session: $F_{(3, 234)} = 0.80$, $p = 0.49$, $\eta^2_p = 0.01$).

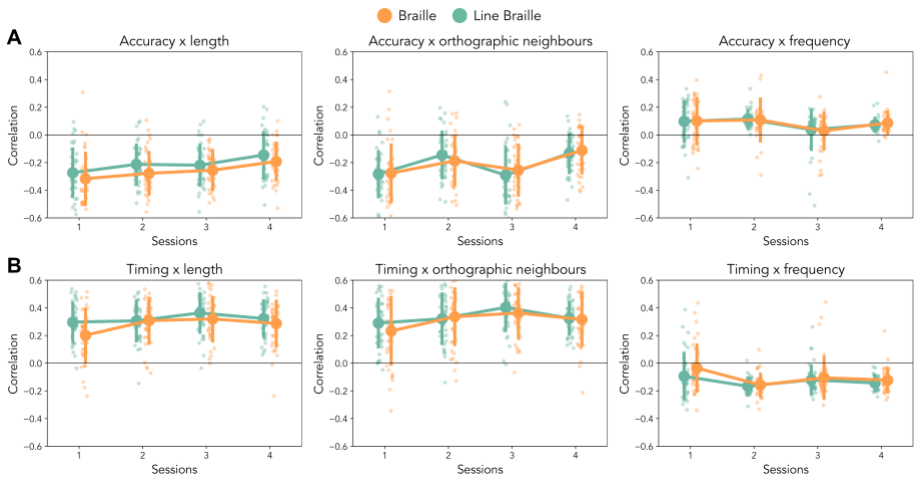

Figure 4 – Correlations between behavioural variables and language statistics. (A) Correlations between accuracy during test and stimulus length (left) show more sensitivity in transcription from Braille than from Line Braille, indicating a lower assimilation of the linguistic components. Words' orthographic neighbours (center), and frequency (right) do not show differences between groups. All variables show a session effect which highlights lower errors with time. (B)

Correlations between writing time during test and stimulus length (left), words' orthographic neighbours (center), and frequency (right) all show only main effects of the session and no group differences. Error bars represent standard error.

Transcription accuracy and reading timing in word training

To further explore possible differences between the two scripts, we measured the accuracy transcription, and reading time during the training on words presented in sessions two, three, and four (Figure 4). For transcription accuracy, we measured accuracy on the 10% of training trials that required a written answer (N = 20). We did not find any effect of script ($F_{(1, 78)} = 0.15$, $p = 0.69$, $\eta^2_p = 0.001$), session ($F_{(3, 156)} = 2.275$, $p = 0.11$, $\eta^2_p = 0.03$), nor interaction ($F_{(3, 156)} = 0.165$, $p = 0.85$, $\eta^2_p = 0.002$). This lack of effect might be caused by the limited number of items tested in this phase, combined with the ceiling performance. Reading time was defined as the amount of time passed between the start of the trial and the time participants pressed the key to visualize the correct transliteration, and could be assessed on the remaining 180 trials for each training session. In this case, we observe a main effect of session ($F_{(3, 156)} = 81.347$, $p < 0.001$, $\eta^2_p = 051$) associated to the progression during the whole experiment. Moreover, we do not observe any significant difference between scripts ($F_{(1, 78)} = 0.975$, $p = 0.33$, $\eta^2_p = 0.01$) and a marginally significant interaction ($F_{(3, 156)} = 2.967$, $p = 0.054$, $\eta^2_p = 0.04$). The trend of the interaction, albeit not significant, is consistent with the interaction found on accuracy during testing, with a tendency towards a difference between scripts early in training (here: slower reading time for Braille) and no sign of any difference in the last session. More in depth, the post-hoc t-tests do not highlight any significant script difference during any session (session 2: $t_{(78)} = 1.66$, $p_{uncorr} = 0.10$, $p_{Bonf} = 0.31$, BF = 0.76, err. < 0.001 ; session 3: $t_{(78)} = 0.09$, $p_{uncorr} = 0.93$,

p_{Bonf} = 1, BF = 0.23, err. < 0.001; session 4: $t_{(78)}$ = 0.70, p_{uncorr} = 0.48, p_{Bonf} = 1, BF = 0.29, err. < 0.001).

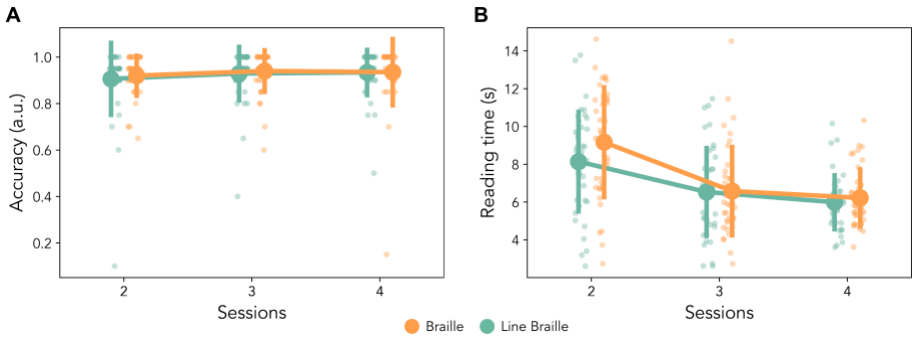

Figure 4 – Accuracy and reading time during training. (A) Accuracy in the transcriptions during the training. No significant result emerged from the analysis, the performance is already at ceiling in both groups from the first training session on words. (B) Time spent reading the words in the novel alphabet. Participants need less time to process the word when training progresses. Error bars represent standard error.

Discussion

We conducted a four-days online behavioural training where participants learnt to read either visual Braille or a custom script with dots connected into lines. Overall, we showed that the accuracy at a transcription test improves over time in a similar fashion between scripts. There was a small advantage for the line Braille in the speed of learning, reflected in a significant interaction between script and session. Writing time during test improved over learning, but did not differ between scripts. Linguistic factors, including word length, number of orthographic neighbours, and frequency, affected task accuracy and writing time, but mostly did so in a similar way in the two scripts. Overall, we show comparable learning trajectories between Braille and a line-based

script, hinting at a limited role of the presence of lines and junctions in scripts, except for a small and temporary effect of the presence/absence of lines on the speed of learning.

The interaction shows that transcription from line Braille is more accurate than from Braille with dots earlier in training, revealing at best a temporary advantage of the presence of lines. Additionally, we also see an overall stronger correlation between stimulus length and Braille script compared to the line-based one. Such higher sensibility to length could also be a proxy of a slightly weaker mapping between the novel and the native alphabet. Together, this evidence could lead to the conclusion that some line advantage is indeed present. After all, the two scripts are visually very different, and no visual script evolved naturally to be represented by dots (Changizi et al., 2006). It remains possible that transcriptions from Braille are more prone to errors, reflecting a less strong mapping. However, the differences between scripts are very temporary and towards the end of four days of training there is no difference in accuracy during test, nor in reading speed during training.

The conclusion that there are no differences between the scripts towards the end of training is further strengthened by the Bayes Factors, which show evidence in favour of the null hypothesis in the last training session. We measured similar learning accuracy of the novel scripts, regardless its presence or absence of line-based visual features. Previous literature hypothesised an advantage of line junctions processing when reading that builds over the already present circuitry for object recognition (Dehaene et al., 2005; Dehaene & Cohen, 2011). The importance of line junctions as building blocks of a visual

script suggested hindered processing when such features were absent (Szwed et al., 2009, 2011). Moreover, while many conducted training experiments on peculiar scripts, derived from example from other categories found to activate areas in vOTC, like houses and faces (Martin et al., 2019; Moore et al., 2014), those studies looked at more complex stimuli that potentially recruit the same organizational processes of line junctions integration. Rarely there has been a direct comparison between scripts with different organizational principles, line junctions integration versus dots patterns. Bola and colleagues (Bola et al., 2017) did so and compared learning visual Braille to Cyrillic. The authors show how learning Cyrillic, being made out of canonical line junctions, poses an advantage to achieve fluent reading. However, differences remained between training programs. In our study, we balanced such conditions: participants undergo the same training regime, with consistent timings and sessions across groups; participants see equally novel stimuli that rarely have been seen outside the experimental setting and the same transcribed words in their native script (Latin). We show that when the trainings are balanced, the advantage of line junctions is only very temporary and disappears with training. We also explored whether one script would be more sensitive to the linguistic properties, namely stimulus length, word frequency, and orthographic neighbours of the stimuli presented. These linguistic factors influenced task performance in obvious ways, with for example correlations between accuracy and writing time on the one hand and word length and the number of orthographic neighbours on the other hand. However, the effects of these factors almost never differed between the two scripts, corroborating the overall similarities in learning a new alphabet independently of the presence of lines and line junctions.

One limitation of our testing protocol is that the stimuli were always presented for 6 seconds. With this protocol we might miss effects in speed of reading. Yet this does not seem to be the case if we take the results as a whole. First, maximum accuracy at test is 80%, which is lower than the ceiling and lower than accuracy during training when there is no time limit (the latter shown in Figure 4A). This suggests that presentation time during the test phase is brief enough to avoid ceiling effects, and warrants evaluating group differences. Second, even during training, where presentation duration is self-paced, there are no indications of differences in the processing speed during the later sessions (Figure 4B). Overall, and also considering the reported Bayes Factors, our findings indicate that the Braille with dots and the Line Braille with connected dots are associated with equal accuracy and speed in the later training and test sessions.

The proficiency of participants to learn the visual Braille with dots and the effect of factors like orthographic neighbours are consistent with our previous observations that expertise in visual Braille engages the same reading network and neural coding principles as line-based scripts. We previously observed that extensive training in visual Braille leads to similar activity across the reading brain network when compared to the ones observed in the participants' native Latin alphabet (Cerpelloni et al., 2024). Given the malleability of the reading network to adapt to various scripts (Baker et al., 2007; Martin et al., 2019; Moore et al., 2014; Xue & Poldrack, 2007) and in particular to tactile and visual Braille (Cerpelloni et al., 2024; Reich et al., 2011; Siuda-Krzywicka et al., 2016),

we hypothesize that our training procedure rely on such adaptability to trigger equal performance across script types.

All this evidence could be seen in favour of concurrent theories about reading in the brain, which stress the importance of the linguistic content of the stimuli rather than their visual features. Under this assumption, reading could be seen as a mapping of symbols to meaning, a mapping not restricted by the visual features of the stimuli.

In conclusion, we show that the presence or absence of line junctions in a novel script, under balanced training conditions, is only associated with a temporary difference in the speed of learning but no differences in later training phases, illustrating the flexibility of the human visual system and reading network to adapt to a wide range script structure.

References

Baker, C. I., Liu, J., Wald, L. L., Kwong, K. K., Benner, T., & Kanwisher, N. (2007). Visual word processing and experiential origins of functional selectivity in human extrastriate cortex. *Proceedings of the National Academy of Sciences*, *104*(21), 9087–9092. https://doi.org/10.1073/pnas.0703300104

Bola, Ł., Radziun, D., Siuda-Krzywicka, K., Sowa, J. E., Paplińska, M., Sumera, E., & Szwed, M. (2017). Universal Visual Features Might Be Necessary for Fluent Reading. A Longitudinal Study of Visual Reading in Braille and

Cyrillic Alphabets. *Frontiers in Psychology*, *8*. https://doi.org/10.3389/fpsyg.2017.00514

Brysbaert, M., Stevens, M., Mandera, P., & Keuleers, E. (2016). The impact of word prevalence on lexical decision times: Evidence from the Dutch Lexicon Project 2. *Journal of Experimental Psychology: Human Perception and Performance*, *42*(3), 441–458. https://doi.org/10.1037/xhp0000159

Büchel, C., Price, C., & Friston, K. (1998). A multimodal language region in the ventral visual pathway. *Nature*, *394*(6690), 274–277. https://doi.org/10.1038/28389

Cerpelloni, F., Van Audenhaege, A., Matuszewski, J., Gau, R., Battal, C., Falagiarda, F., Op de Beeck, H., & Collignon, O. (2024). Expertise in reading visual Braille co-opts the reading network. *bioRxiv*, 2024.05.08.593104. https://doi.org/10.1101/2024.05.08.593104

Changizi, M. A., Zhang, Q., Ye, H., & Shimojo, S. (2006). *The Structures of Letters and Symbols throughout Human History Are Selected to Match Those Found in Objects in Natural Scenes.* 23.

Cohen, L., Dehaene, S., Naccache, L., Lehéricy, S., Dehaene-Lambertz, G., Hénaff, M.-A., & Michel, F. (2000). The visual word form area: Spatial and temporal characterization of an initial stage of reading in normal subjects and posterior split-brain patients. *Brain: A Journal of Neurology*, *123*, 291–307.

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

Dehaene, S., & Cohen, L. (2011). The unique role of the visual word form area in reading. *Trends in Cognitive Sciences*, *15*(6), 254–262. https://doi.org/10.1016/j.tics.2011.04.003

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

Dehaene-Lambertz, G., Monzalvo, K., & Dehaene, S. (2018). The emergence of the visual word form: Longitudinal evolution of category-specific ventral visual areas during reading acquisition. *PLOS Biology*, *16*(3), e2004103. https://doi.org/10.1371/journal.pbio.2004103

Gaca, M., Olszewska, A. M., Droździel, D., Kulesza, A., Kossowski, B., Jednoróg, K., Matuszewski, J., & Marchewka, A. (n.d.). *How learning to read Braille in visual and tactile domains reorganizes the sighted brain*.

Gastoud, E. (2022). *Nécessité des caractéristiques visuelles universelles dans l'efficacité de la lecture: Le cas du braille visuel* [PhD Thesis, UCL - Faculté de psychologie et des sciences de l'éducation]. http://hdl.handle.net/2078.1/thesis:34806

Kass, R. E., & Raftery, A. E. (1995). Bayes Factors. *Journal of the American Statistical Association*, *90*(430), 773–795. https://doi.org/10.1080/01621459.1995.10476572

Keuleers, E., & Brysbaert, M. (2010). Wuggy: A multilingual pseudoword generator. *Behavior Research Methods*, *42*(3), 627–633. https://doi.org/10.3758/BRM.42.3.627

Logothetis, N. K., & Sheinberg, D. L. (1996). *Visual Object Recognition*.

Martin, L., Durisko, C., Moore, M. W., Coutanche, M. N., Chen, D., & Fiez, J. A. (2019). The VWFA Is the Home of Orthographic Learning When Houses

Are Used as Letters. *Eneuro*, *6*(1), ENEURO.0425-17.2019. https://doi.org/10.1523/ENEURO.0425-17.2019

Moore, M. W., Durisko, C., Perfetti, C. A., & Fiez, J. A. (2014). Learning to Read an Alphabet of Human Faces Produces Left-lateralized Training Effects in the Fusiform Gyrus. *Journal of Cognitive Neuroscience*, *26*(4), 896–913. https://doi.org/10.1162/jocn_a_00506

Peirce, J. W. (2007). PsychoPy—Psychophysics software in Python. *Journal of Neuroscience Methods*, *162*(1–2), 8–13. https://doi.org/10.1016/j.jneumeth.2006.11.017

Peirce, J. W. (2008). Generating stimuli for neuroscience using PsychoPy. *Frontiers in Neuroinformatics*, *2*. https://doi.org/10.3389/neuro.11.010.2008

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

Riesenhuber, M., & Poggio, T. (1999). Hierarchical models of object recognition in cortex. *Nature Neuroscience*, *2*(11), 1019–1025. https://doi.org/10.1038/14819

Riesenhuber, M., & Poggio, T. (2002). Neural mechanisms of object recognition. *Current Opinion in Neurobiology*, *12*(2), 162–168. https://doi.org/10.1016/S0959-4388(02)00304-5

Saygin, Z. M., Osher, D. E., Norton, E. S., Youssoufian, D. A., Beach, S. D., Feather, J., Gaab, N., Gabrieli, J. D. E., & Kanwisher, N. (2016). Connectivity precedes function in the development of the visual word form area. *Nature Neuroscience*, *19*(9), 1250–1255. https://doi.org/10.1038/nn.4354

Siuda-Krzywicka, K., Bola, Ł., Paplińska, M., Sumera, E., Jednoróg, K., Marchewka, A., Śliwińska, M. W., Amedi, A., & Szwed, M. (2016). Massive cortical reorganization in sighted Braille readers. *eLife*, *5*, e10762. https://doi.org/10.7554/eLife.10762

Szwed, M., Cohen, L., Qiao, E., & Dehaene, S. (2009). The role of invariant line junctions in object and visual word recognition. *Vision Research*, *49*(7), 718–725. https://doi.org/10.1016/j.visres.2009.01.003

Szwed, M., Dehaene, S., Kleinschmidt, A., Eger, E., Valabrègue, R., Amadon, A., & Cohen, L. (2011). Specialization for written words over objects in the visual cortex. *NeuroImage*, *56*(1), 330–344. https://doi.org/10.1016/j.neuroimage.2011.01.073

Vinckier, F., Dehaene, S., Jobert, A., Dubus, J. P., Sigman, M., & Cohen, L. (2007). Hierarchical Coding of Letter Strings in the Ventral Stream: Dissecting the Inner Organization of the Visual Word-Form System. *Neuron*, *55*(1), 143–156. https://doi.org/10.1016/j.neuron.2007.05.031

Xue, G., & Poldrack, R. A. (2007). The Neural Substrates of Visual Perceptual Learning of Words: Implications for the Visual Word Form Area Hypothesis. *Journal of Cognitive Neuroscience*, *19*(10), 1643–1655. https://doi.org/10.1162/jocn.2007.19.10.1643

Declaration statements

Funding

The project was funded in parts by a Mandat d'Impulsion Scientifique awarded to OC, the Belgian Excellence of Science (EOS) program (Project No. 30991544) awarded to HO and OC, a Flagship ERA-NET grant SoundSight (FRS-FNRS PINT-MULTI R.8008.19) awarded to OC, and research project G0D3322N of the Fund for Scientific Research (FWO) Flanders and Methusalem project METH/24/003 awarded to HO. OC is a senior research associate at the Fond National de la Recherche Scientifique de Belgique (FRS-FNRS). FC received funding from a KU Leuven-UCLouvain cooperation grant.

Conflicts of Interest / Competing interests

The authors declare no conflicts of interest and no competing interests.

Ethics Approval

The research has been approved by the Sociaal Maatschappelijke Ethische Commissie (Social and Societal Ethics Committee, SMEC) of KU Leuven (Reference: G-2020-1902)

Consent to participate

All participants gave explicit consent to the treatment of their data at the start of the first experiment session, where they were prompted to read and accept the informed consent displayed on their screen.

Consent for publication

All participants gave explicit consent to publication as part of the informed consent displayed on their screen at the start of the first experiment session.

Availability of data and materials

Data and materials are available on GitHub: experiments (https://github.com/fcerpe/VBT_experiments) and analyses (https://github.com/fcerpe/VBT_data).

Code availability

All code to administer the different training sessions (https://github.com/fcerpe/VBT_experiments) and to analyse the results (https://github.com/fcerpe/VBT_data) is publicly available on GitHub.

Authors' contributions

Author's contributions are detailed in Figure 5.

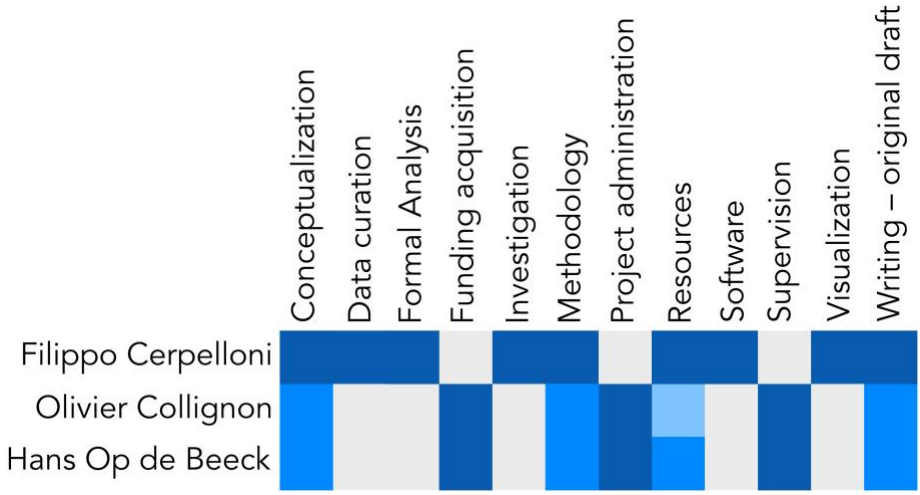

Figure 5 – Author contributions. Based on the CRedIT (https://credit.niso.org/) taxonomy. Each contribution was assessed as 'lead' (dark blue), 'equal' (mid blue), 'support' (light blue).

Chapter 3. How learning to read visual Braille co-opts the reading brain

This chapter as provided in the thesis has been published in the journal "Journal of Cognitive Neuroscience".

Reference: Filippo Cerpelloni, Alice Van Audenhaege, Jacek Matuszewski, Remi Gau, Ceren Battal, Federica Falagiarda, Hans Op de Beeck*, Olivier Collignon*; How Learning to Read Visual Braille Co-opts the Reading Brain. *J Cogn Neurosci* 2025; doi: https://doi.org/10.1162/jocn_a_02341

Abstract

Learning to read assigns linguistic value to an abstract visual code. Whether regions of the reading network tune to visual properties common to most scripts or code for more abstracted units of language remains debated. Here we investigate this question using visual Braille, a script developed for touch that does not share the typical explicit shape information of other alphabets, yet maps onto the same phonology and lexicon as other more regular scripts. First, we compared univariate responses in visual Braille readers and a naïve control group and found that individually localized Visual Word Form Area (VWFA) was selectively activated for visual Braille when compared to scrambled Braille only in expert Braille readers. Multivariate analyses showed that linguistic properties can be decoded from Latin script in both groups and from Braille script in expert readers in an extended network of brain regions including the early visual cortex (V1), the lateral occipital region (LO), the VWFA and the left Posterior Temporal area (l-PosTemp). These results suggest that the tuning of an extended reading network to orthography relies more on the linguistic content of the script rather than their specific visual features (e.g. line junctions). Nevertheless, cross-scripts generalization was significantly lower than within-script decoding and failed to reveal common representations across Latin and Braille in experts in all regions except the l-PosTemp. These results suggest that V1, LO and VWFA encode orthographic representations in a script-specific manner, whereas l-PosTemp encodes abstracted linguistic information.

Introduction

In vision, object recognition is modelled as a hierarchical process going from the detection of simple features to their progressive integration into more complex shapes (Riesenhuber & Poggio, 1999; Biederman, 1987; Krizhevsky et al., 2017). Shape, an arrangement formed by lines joined together in a particular way or by lines around the outer edge of an object, plays therefore a prominent role in human object recognition (Grill-Spector & Malach, 2004; Landau et al., 1988; Tarr & Bulthoff, 1998). The flexibility of shape processing in the human brain has allowed the development of written scripts, highly complex visual stimuli with symbolic meaning. Most written systems developed throughout history share basic shape features with natural objects (Changizi et al., 2006). Core features of visual object recognition may have been co-opted for written word recognition (Dehaene et al., 2005), exerting pressure for scripts to appear the way they do. Supporting this idea, basic sensitivity to Latin script letters seems to be present in untrained monkeys (Rajalingham et al., 2020), further illustrating how culturally invented scripts may have been optimized from the natural scaffolding of shape processing in the brains of primates.

In the human brain, visual scripts activate an extended network of brain regions when compared to non-linguistic visual stimuli (Fedorenko et al., 2011). It has been suggested that a predisposed sensitivity to shape features (e.g. line junctions) represents a proto-organization on which orthographic selectivity develops (Arcaro & Livingstone, 2017; Dehaene et al., 2005). The so called Visual Word Form Area (VWFA; Cohen et al., 2000, 2002), the major focus of

word form selectivity in visual cortex, is a paradigmatic example of this debate. In terms of visual features, the presence of pre-existing shape selectivity in VWFA is one of the most often mentioned causes of the location of this region (Dehaene & Cohen, 2011; Dehaene-Lambertz et al., 2018). The importance of line junctions for fluent reading (Bola et al., 2017), the partial overlap between VWFA and ventral visual areas responding selectively to line junctions (Szwed et al., 2011), equal responses to line drawings of objects and false fonts (Ben-Shachar et al., 2007), and comparable levels of activation for known and foreign characters (Vogel et al., 2012; Xue & Poldrack, 2007) all lead to the suggestion that VWFA, along with other regions located in the fusiform gyrus, might be involved in specific shape processing more generally (Roberts et al., 2013). Even without reading experience, the pre-existing shape selectivity in occipitotemporal cortex might be very well suited for learning to read, as demonstrated by the presence of basic sensitivity to Latin script letters in untrained monkeys (Rajalingham et al., 2020).

However, an alternative account of VWFA downplays the role of visual features and emphasizes the domain- and task-specific role of this region. It regards the integration of visual regions selective to orthography with downstream language regions involved in phonological and semantic processing as the key organisational feature driving selectivity for written language (Dębska et al., 2023; Price & Devlin, 2003, 2004, 2011). Structural and functional connectivity between VWFA and the traditional language network, even before infants learn to read may back up an interactive account of VWFA (Saygin et al., 2016; Li et al., 2020; Stevens et al., 2017). In apparent support of this view, studies have shown activations of VWFA for speech processing (Pattamadilok et al.,

2019; Planton et al., 2019), but this has often been interpreted as an automatic activation of orthographic representations (Cohen et al., 2000, 2002). Further, some studies have shown that learning to read Korean Hangul or a script made of pictures of houses or faces (L. Martin et al., 2019; Moore et al., 2014; Xue & Poldrack, 2007) activates the reading network. However, those "atypical" scripts still rely on the processing of line junctions and shape, potentially triggering their recruitment of the typical reading network. Studies in in sighted and blind people with tactile Braille reading further support the idea that VWFA encodes linguistic information more generally, rather than being selective to specific visual features (Büchel et al., 1998; Reich et al., 2011; Siuda-Krzywicka et al., 2016).

To the extent that brain preference for visual features or linguistic properties each have a separate role to play in the emergence of orthographic selectivity due to experience, the question emerges on how these factors might work together, and what impact each property might have. The findings with house- and face-based scripts and tactile Braille are consistent with the suggestion that the domain-specific task-related connectivity between visual and linguistic regions might determine in which precise region in the visual system a preference for the experienced stimuli will emerge (Saygin et al., 2016). Yet, even if true, this does not exclude a role for visual properties. For example, visual properties might determine which neural populations are implicated within the activated region. So, whereas a regular shape-based script, a fabricated face-based script, and even tactile braille might all activate the reading network, the way these scripts are represented in these brain regions might be different due to their visual properties. Such investigation requires

an in-depth comparison of multiple scripts with methods that have the sensitivity to pick up fine-scale differences in activity patterns. A relevant study was published recently by (Zhan et al., 2023), who showed with high-resolution 7T imaging that the English and Chinese script both activate nearby and sometimes even overlapping brain regions in English/Chinese bilinguals, yet some small patches seem uniquely activated by only one of the two scripts. However, English and Chinese are both line-based scripts, and also differ in their mapping onto phonology and lexical entries - leaving open the possibility that nonvisual linguistic factors might explain the partial separation of the two scripts.

Overall, the literature shows that both visual and domain-specific nonvisual factors influence the activity of VWFA. An assessment of how they might work together in determining how orthographic selectivity emerges due to experience is missing, and visual Braille represents a unique opportunity to assess the contribution of visual and linguistic aspects. It is a visual script that does not share the most basic features of other scripts like lines, junctions, and clear outer shapes thought to be necessary for efficient reading (Bola et al., 2017) and for the location of the reading network. The letters are simple configurations of 1 to 6 dots, and arguably simpler than any other stimulus that is typically used to investigate cortical vision. Yet visual Braille maps onto the same phonology and lexicon as other scripts in the same language community, specifically the Latin script in a western country. Therefore, visual Braille provides a strong difference in visual properties but an optimal matching of nonvisual linguistic properties.

In the present study, we used fMRI to individually localize VWFA and to precisely evaluate whether the word form system activated for a Latin script is co-activated by reading visual Braille (Figure 1A). We then performed multivariate pattern analysis (MVPA) on VWFA, on several other key areas involved in visual processing, early visual cortex (V1) and lateral Occipital (LO), and on left Posterior Temporal area in the language network. Through MVPA we tested whether the same coding principles (using Representational Similarity Analyses, RSA) and neural populations (using cross-script decoding) underly the processing of both scripts (Figure 1B). We observed that most regions tested present overwhelming similarities in the coding principles used to processing Latin and Braille scripts, showing the flexibility of the visual system and the reading network for scripts made of atypical visual features like Braille. Importantly however, in most of the regions including VWFA, these coding principles were implemented in a script-specific manner showing intermingled but separate representations of both scripts. Together, these findings reveal how the word processing system implements the same function of reading applied on visual input with fundamentally different properties.

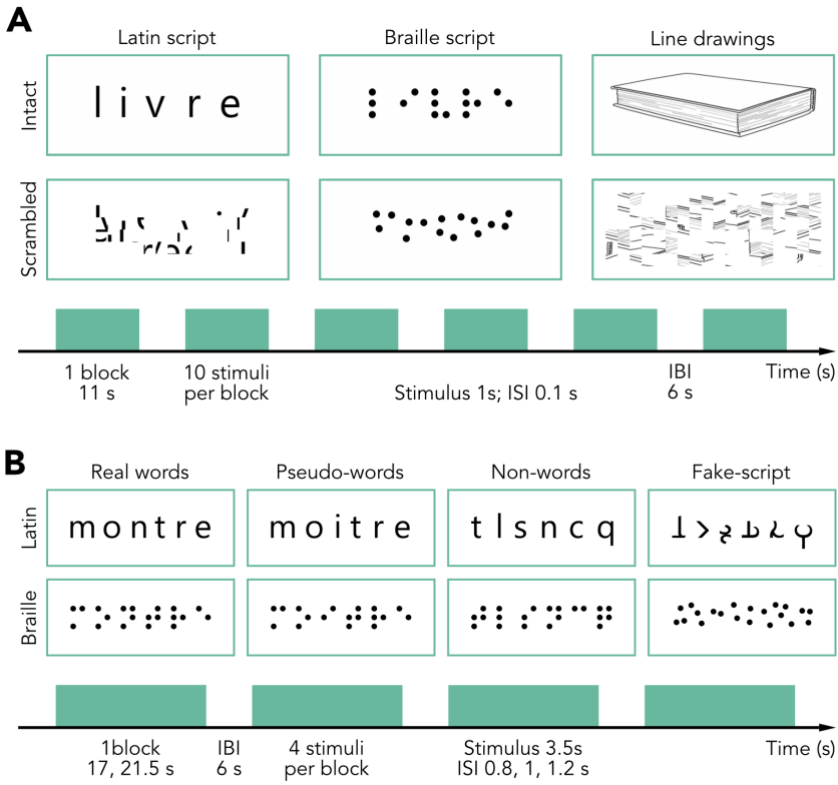

Figure 5 — Experimental design and examples of stimuli for experiments 1 and 2. (A) Experiment 1: functional localizer. We presented six categories of stimuli (intact and scrambled Latin-script words, Braille words and line drawings) to participants in an fMRI block design. Each participant underwent two runs in which six blocks for each category were presented in pseudo-randomized order. Blocks lasted for 11 seconds and were separated by a fixed inter-block interval (IBI) of 6 seconds. Each block consisted of ten stimuli from one category, each for a duration of 1 second and with an inter-stimulus interval (ISI) of 0.1s. During the session, participants performed a 1-back task to detect the repetition of a stimulus. The target was present in 10% of the blocks. **(B)** Experiment 2: Coding of linguistic properties. Four different linguistic conditions (Real Words, Pseudo-Words, Non-Words, Fake Script) for each script (Latin, Braille) were used. We presented these categories of stimuli (intact and scrambled Latin-script words, Braille words, line drawings) to participants in a fMRI block design. Each participant underwent twelve sessions in total, six for each script (counterbalanced across subjects). In each run, three blocks for each category of one script were presented in a pseudo-randomized order. Each block lasted for 17 or 21.5 seconds, depending on the presence of a target, with a 6-second IBI separating the two blocks. Four stimuli from the same category were presented within a block, each for 3.5 seconds, with a jittered ISI of 0.8, 1, or 1.2 seconds. During the session, participants performed a 1-back task to detect the repetition of a stimulus. The target was present in 10% of the blocks. The position of the target, as well as jittered ISI, block order, and stimuli order, were counterbalanced across runs.

Materials and Methods

Participants

A total of twenty-three participants took part in the data acquisition. Experts in visual Braille reading are rare, so we knew a priori that the expert group could be small with number of participants mostly determined by availability. To overcome problems due to statistical power, we decided to recruit twice as many control participants. Furthermore, we performed extensive piloting in other participants to select two paradigms (localizer and multivariate imaging) that provide interpretable results in the main region of interest (visual word form area) in individual participants.

Data from five subjects were discarded due to incomplete acquisitions (three control subjects) or inability to functionally localize VWFA for the Latin-based script (two subjects: one control, one expert), resulting in eighteen subjects in the final sample. Six of them (6 females, mean age = 38.2 ± 11.4) were experts in reading Braille visually and twelve (10 females, mean age = 39.3 ± 10.4) were controls naïve to Braille alphabet. To determine visual Braille expertise, we asked expert participants to report their use of visual Braille as a declarative number of years spent working with / using Braille (mean expertise = 5.25 ± 5.43 years). Furthermore, we assessed their reading speed through a behavioural reading test. All expert readers were presented with the same list of 30 words and instructed to read as many words as possible out loud in the given time of 1 minute (mean number of words read = 12.2 ± 6.11 words; range = 6-23).

Experimental design

Each participant took part in one experimental session, lasting maximum two hours in total, of which maximum 90 minutes inside the MRI scanner. Four participants did not complete the whole experiment in one session due to time constraints, and were invited to a second, shorter, session to conclude data acquisition. MRI acquisition followed the same procedure for all the participants. In the scanner, we acquired a structural image followed by two functional scans of the localizer experiment and twelve functional scans of the MVPA-oriented experiment, alternating between one run with Latin script stimuli and one with Braille script stimuli. The stimuli and the counterbalancing procedures used in each experiment are detailed below. All stimuli were presented in white on a black background, on a screen with 1920x1200 pixels of resolution, placed at 170 cm and viewed by the participant through a mirror placed on the headset.

Functional localization experiment

We designed a functional localizer to individually define VWFA, selective for words, and bilateral Lateral Occipital area (LO), selective for object shape. We included six conditions of stimuli (Figure 1A): concrete French words expressed in the Latin alphabet (FW), the same French words presented in the Braille alphabet (BW), line drawings of the objects represented by the words (LD), and scrambled conditions for each of the 'intact' categories (scrambled Latin words, SFW; scrambled Braille words, SBW; scrambled line drawings, SLD). We selected 20 items for each condition. Latin words were selected from the Lexique 3.83 (New et al., 2005) and comprised of a length between 4 and 8 letters, and a frequency between 1 and 20 appearances per million words.

Braille script words were the literal transcription of the Latin-based script words (*Code Braille Français Uniformisé Pour La Transcription Des Textes Imprimés (CBFU)*, 2008). Line drawings came from two images datasets (Brodeur et al., 2010; Rossion & Pourtois, 2004). If necessary, we manipulated the colour information by obtaining grayscale stimuli that presented only the essential lines of the drawings. Scrambled conditions were created through two different custom algorithms, applied to each 'intact' stimulus. Scrambling of Latin words and line drawings were the result of the re-arrangement of different parts of the corresponding 'intact' stimulus image in a grid-like manner. The scrambled Braille condition was created by randomizing the position of dots of each 'intact' Braille word. We used different scrambling methods to ensure the alteration of each word's visual properties, line junctions in the case of Latin words and organization of dots in the case of Braille words. The localizer experiment included two block-design runs. In each run, participants were presented with 6 blocks of 10 stimuli presented for 1 second in white on a black background (Inter Stimulus Interval, ISI = 0.1s; Inter Block Interval, IBI = 6s). The order of the blocks was randomized, while ensuring that one condition was not presented in two consecutive blocks and that the first three blocks of a run were also presented at the end of the same run, in reverse order. Participants performed a 1-back task with a button press and were asked to detect an immediate repetition of the same stimulus (10% of the trials, so once per block). All stimuli subtended a frame of 520 x 208 pixels, corresponding to 6.1 and 2.5 degrees of visual angle in the MRI scanner, respectively.

MVPA-oriented experiment

Inspired by Vinckier and colleagues (Vinckier et al., 2007), we designed four different conditions per script in a linguistic gradient fashion (Figure 1B). (1) Real Words (RW) possessed semantic, phonological, and orthographical properties. This set was composed of French words (New et al., 2005) six letters length, with high frequency (>1 per million), and high identifiability (definition of lemma > 90%). Additionally, we controlled bigram and trigram frequency (top 45th percentile; Gimenes et al., 2020). (2) Pseudo-Words (PW) were based on Real Words by identifying stimuli from the French Lexicon Project (Ferrand et al., 2010). We ensured bi- and tri-grams composing Pseudo-Words to have the same frequency of the real word set. Additionally, we measured Levensthein distance (the number of additions, deletions, transformations necessary to transform one string into another) and only included elements with a distance of 2, to avoid excessive similarity and possible confusion with the subset of Real Words. The final set of Pseudo-Word stimuli possesses both orthographic and phonological regularities similar to Real Words, but with limited semantic meaning, and without a lexical entry. (3) For Non-Words (NW), we selected strings of consonants based on infrequent bigrams and trigrams (Gimenes et al., 2020), to create phonological distance between sets. We created the Braille conditions for all these categories by transliteration of the Latin conditions. (4) Fake Script (FS). For the Latin Fake Script, we transliterated the Non-Words set to a custom-made alphabet that maintained the same features of the original consonants, but in re-arranged line junctions. For Braille Fake Script, given the lack of evident shapes to re-arrange, we scrambled the dots arrangement as for the functional localizer.

For each condition, we selected twelve six-characters long different stimuli. This composition of stimuli, comparable with the one made by Vinckier and colleagues (Vinckier et al., 2007), created a linguistic gradient of words with meaning, phonology, and orthography, progressively stripped of one of those properties. The experiment was divided into twelve runs of blocked design, alternating between runs with Latin and with Braille stimuli (Figure 1B). Participants performed a 1-back task, reacting to stimuli repetitions (targets). The presentation of stimuli within a block and run was pseudo-randomized to balance the appearance of targets across runs, categories, and position of the stimulus repeated. A total of six targets in each run were presented (12.5% of all stimuli). We controlled for the order of blocks and, within a block, for the order of categories and single stimuli. All parameters were balanced, resulting in a fixed experiment design which participants performed almost in the same order, aside from the counter-balanced alternation of the starting script, either Latin or Braille. During each block, four stimuli were presented in white on a black background for 3.5s (ISI jittered = 0.8s, 1s, 1.2s; IBI = 6s), each block lasting ~ 17-21.5s including target. The long stimulus presentation time is based on results from a previous study on experts' habituation to Braille words (Coquillart et al., 2022). All stimuli subtended a frame of 500 x 200 pixels, corresponding to 5.9 and 2.4 degrees of visual angle in the MRI scanner, respectively.

fMRI acquisition and pre-processing

This method section was automatically generated using bidspm (Gau et al., 2023). We tested all the participants using a 3T GE (Signa™ Premier, General Electric Company, USA, serial number: 000000210036MR03, software: 28\LX\MR, Software release: RX28.0_R04_2020.b) MRI scanner with a 48-channel head coil at the University Hospital of St Luc, Belgium. We acquired a whole-brain T1-weighted anatomical scan (3D-MPRAGE; 1.0 x 1.0 mm in-plane resolution; slice thickness 1mm; no gap; inversion time = 900 ms; repetition time (TR) = 2189,12 ms; echo time (TE) = 2.96 ms; flip angle (FA) = 8°; 156 slices; field of view (FOV) = 256 x 256 mm2; matrix size= 256 X 256). All functional runs were T2*-weighted scans acquired with Echo Planar and gradient recalled (EP/GR) imaging (2.6 x 2.6 mm in-plane resolution; slice thickness = 2.6mm; no gap; Multiband factor = 2; TR = 1750ms, TE = 30ms, 58 interleaved ascending slices; FA = 75°; FOV = 220 x 220 mm2; matrix size = 84 X 84). We extracted NIfTI files using MRIcroGL (https://github.com/rordenlab/MRIcroGL) and converted the content in BIDS format (Gorgolewski et al., 2016). We pre-processed the (f)MRI acquisitions with bidspm (Gau et al., 2023) (v3.1.0dev) using statistical parametric mapping SPM12 (version 7771) using MATLAB (R2021b) on a macOS computer (13.0.1). The pre-processing consisted of slice timing correction, realignment and unwarping, segmentation and skull-stripping, normalization to the Montreal Neuroimaging Institute (MNI) standard space, smoothing. Slice timing correction used the 28th slice as a reference (interpolation: sinc interpolation). Functional scans were realigned and unwarped using the mean image as a reference (SPM single pass with 6 degrees of freedom). The anatomical image was corrected for bias and segmented using a unified segmentation. The mean functional image from

realignment was co-registered to the bias-corrected anatomical image (degrees of freedom: 6). The transformation matrix from this co-registration was applied to all the functional images. The deformation field was applied to all the functional images (target space: IXI549Space; interpolation: 4th-degree b-spline). Finally, we applied spatial smoothing using a 3D Gaussian kernel (functional localizer, FWHM = 6 mm; main experiment, FWHM = 2mm). These smoothed, normalised BOLD images served as inputs to statistical analyses.

General Linear Models (GLMs)

For both experiments, we analysed fMRI acquisitions with bidspm (Gau et al., 2023) (v3.1.0dev) and SPM12 (version 7771) using the same MATLAB (R2021b) on a macOS computer (13.0.1). For the localizer experiment, at the subject level, we performed a univariate analysis with a linear regression at each voxel of the brain, using generalized least squares with a global FAST model and a high-pass filter (278 seconds cut-off). Image intensity scaling was done run-wide before statistical modelling such that the mean image would have a mean intracerebral intensity of 100. Block timings were convolved with SPM canonical hemodynamic response function (HRF) for the 6 stimuli categories (intact and scrambled Latin words, Braille words, Line drawings), in addition to 'response' and 'target' events of no interest and 6 head motion parameters (three translations, on the X, Y, Z axes, and three rotations) to account for motion artefacts. We computed four main contrasts: intact over scrambled Latin words (FW > SFW), Braille words over scrambled dots (BW > SBW), intact over scrambled line drawings (LD > SLD), intact and scrambled Latin words over rest (FW+SFW > rest). We used the [FW > SFW] and [BW > SBW] contrasts to independently localize VWFA in all the subjects, for Latin and for Braille. We

tested the statistical significance of clusters identified in the [BW > SBW] contrast in or near the Latin-based VWFA, by applying Small Volume Correction (SVC) within a 10mm radius from the peak coordinates extracted from [FW > SFW]. We localized bilaterally the Lateral Occipital area (LO) by the [LD > SLD] contrast. Lastly, we localized V1 by the [FW+SFW > rest] contrast (see *Definition of the Regions of Interest* below). In the MVPA-oriented experiment GLM, we included 8 regressors of interest for stimuli conditions (Real Words, Pseudo-Words, Non-Words, Fake Script for both Latin and Braille scripts), and 8 regressors of no-interest (1 for targets; 1 for responses; 6 for head motion parameters). Contrast images were spatially smoothed using a 3D Gaussian kernel (FWHM = 2). The results of the main GLM were concatenated into 4D maps, to be included in the multivariate pattern analyses.

Eye-movements control analysis

To exclude potential confounds of eye movements, we computed eye displacement and integrated it in a General Linear Model (GLM) as a regressor of no interest. We used bidsMReye (Gau & Cabee, 2023), a BIDS app relying on deepMReye (Frey et al., 2021) to decode eye motion from fMRI time series data, in particular from the MR signal of the eyeballs. Each run, the following values were computed: variance for the X gaze position, variance for the Y gaze position, framewise gaze displacement, number of X gaze position outliers, number of Y gaze position outliers, and number of gaze displacement outliers. Outliers were robustly estimated using an implementation of the median rule (Carling, 2000). We then used the newly assessed parameters to perform two additional subject-level GLM analyses. We added the eye displacement for each run to the regressors of no interest, and computed the same contrasts of

the localizer GLM analysis. This analysis showed very similar findings as the main analysis without these regressors (Figure 2-3). Therefore, the definition of the regions of interest described in the next section refers to the GLM described at the beginning of the section "General Linear Model".

Definition of the Regions of Interest (ROIs)

To investigate the gradient of linguistic representations across low-level sensory and higher-order associative brain regions we defined multiple ROIs, namely: V1, LO, VWFA and left Posterior Temporal cortex (PosTemp), using intersections between neuroanatomy, existing literature and brain activation from our functional localizer. For VWFA, we started from the canonical coordinates (Cohen et al., 2000) of VWFA ([-45, -56, -16]) and we identified the peak activation in each subject corresponding to the [FW > SFW] contrast within a 10mm sphere from the canonical coordinates. We then expanded a progressively larger sphere from the individual peak coordinates and intersected this sphere with the activation map of the same [FW > SFW] contrast until the ROI comprised 100 voxels. Lastly, we constrained the expanded ROI to activation reported in the literature to ensure the precise localization of the area by overlapping the expanded mask with the functional activations downloaded from the neurosynth database (Yarkoni et al., 2011) (term: "visual words"). This procedure permitted us to control for the individual localization of VWFA and to avoid including voxels from adjacent but unrelated areas. To ensure the robustness of our results, we compared it to a classical definition of ROIs from neurosynth.org alone (entry keyword: "visual words"). We extracted the highest-responding 100 voxels from the activation map and performed the same analysis pipeline reported below in relation to

our custom model. As reported in Figure 3-1 in the Appendix, we do not observe differences between the two methods. We continued the analyses with our individual definition of ROIs as we can better ensure the fine spatial localization of the areas (Fedorenko et al., 2010). We used the same approach to localize bilateral LO from the [LD > SLD] contrast. More specifically, we started from canonical coordinates (Left LO: [-47, -70, -5], right LO: [47, -70, -5]) and applied a 10mm sphere around each set of coordinates to identify individual peaks in the localized [LD > SLD] activation. We then expanded a sphere from the individual peaks until a minimum threshold of 100 voxels was reached. Lastly, we overlapped the expansion with the neurosynth mask for the term "object". To identify early visual cortex (V1), we overlapped the bilateral brain mask from the Anatomy toolbox (Eickhoff et al., 2005) for bilateral V1 with a [FW + SFW > rest] contrast. For subsequent analyses on areas of the language network, we relied on parcels from Fedorenko and colleagues (Fedorenko et al., 2010). For each subject, we overlapped activations for the [FW > SFW] contrast to the different language parcels to identify brain regions that consistently responded to linguistic stimuli across subjects. Although in different subjects we could identify many areas of both the left and right hemispheres, we focused our analyses solely on the left Posterior Temporal region (l-PosTemp), as it was the only area to be present reliably in individual subjects, being identified in 15 out of the 18 total participants analysed, corresponding to more than 80% consensus in the activations.

Multivariate Pattern Analyses

To test the gradient of linguistic representations we performed multivariate pattern analyses using the CoSMoMVPA toolbox (Oosterhof et al., 2016) combined with custom MATLAB code (version 2021b). We applied this method to all the regions described in the section "ROI definition". To obtain decoding accuracies, we performed feature selection across individuals by restricting the number of features to correspond to the number of voxels in the smallest ROI extracted through each method. This resulted in the feature selection of 43 voxels for LO and VWFA, extracted through the functional localizers, 62 for l-PosTemp, and 81 for V1. We then trained a support vector machine ('libsvm') in a leave-one-run-out fashion. To identify an organization of linguistic content, we performed pair-wise decoding, obtaining six comparisons per script (RW-PW, RW-NW, RW-FS, PW-NW, PW-FS, NW-FS). To compare the representational patterns emerging from the pairwise decoding accuracies, we performed Representational Similarity Analysis (Kriegeskorte et al., 2008) (RSA). We created a Representational Dissimilarity Matrix (RDM) for each group and script, where the value in each cell is given by the binary decoding accuracy of the classifier for that comparison. Resulting RDMs represent the linguistic (dis)similarity between stimulus conditions. Additionally, we correlated each neural RDM to a model of linguistic distance between categories. We assigned a value to each category of stimuli based on the number of linguistic features possessed. Concretely, Real Words possess orthographic, phonological, and semantic properties (3 properties), Pseudo Words might possess some low semantic meaning but are void of a lexical entry (2 properties), Non Words have orthographic but low frequency (non-pronounceable) phonological elements (1 property), and Fake Script has none. The resulting distances were

computed by subtracting the number of features shared and normalized by dividing for the number of properties of the Real Words. We performed the correlations between neural matrices by applying non-parametrical statistical tests (Stelzer et al., 2013; Mattioni et al., 2020). When comparing two conditions, we first computed the correlation for each individual of the first condition to the group average of the second condition. For conditions from the same group (e.g. experts-Latin and experts-Braille), we averaged the group RDM of the second condition by excluding the subject being compared. We then averaged the individual correlations. Next, we performed the same correlation in the other direction, correlating each individual RDM from the second condition to the first condition averaged (with the same precaution), and then calculated the mean of these correlations. We then averaged the correlations obtained from the two directions, obtaining the correlation between the two matrices. Lastly, to assess the generalizability of the decoding from one script to the other, we performed cross-script decoding. We limited our analyses to visual Braille experts, the only group showing significant decoding for both scripts in our ROIs. For each script, we trained our multi-voxel classifier on pair-wise linguistic differences within a script and tested the classification on the other script (e.g., train on the difference between Braille Real Words and Pseudo-Words and test on the same difference within the Latin script). We performed this classification reversely for both alphabets and averaged across the two directions. To control for possible collinearity between the linguistic and the visual properties of the two scripts, we computed additional models based on the difference in pixels between two stimuli of the same script, comparing each stimulus to every other based on the number of overlapping white pixels between the images. We then

averaged the pixel differences within one condition and one script to obtain two 4-by-4 matrices representing the visual differences across categories. We then used the same procedure described above to compute partial correlations between one neural RDM and one model (linguistic or visual), controlling for the effects of the other model (visual or linguistic). We limited this analysis to VWFA, as it was the main focus of our study. The results are presented in Figure 3-2. We observed that both visual and linguistic factors seems to contribute to the brain representations of scripts in VWFA, and that these factors are modulated by experience, as we do not observe correlations between the brain representations of Braille in the non-experts for neither visual nor linguistic models.

Statistical analyses

Given the inter-subject variability in the location of functional regions, all our analyses were carried out in individually localized ROIs (see ROI definition). To test between groups differences in VWFA activation for the [BW > SBW] contrast, we applied a Chi-square test on the number of participants who showed activation for such contrast. To test statistical significance of overall decoding accuracies we performed one-tail t-tests against chance level (mu = 0.5). To test group-script differences, we performed two-tailed t-tests with other script-group pairs to test for significant differences in the decoding accuracy. If the pairs referred to the same group (e.g. Latin, experts and Braille, experts), we performed paired t-tests. In all the t-tests, we additionally calculated Cohen's D effect size measure. To test whether cross-script decoding was significantly different from within-script decoding, we averaged the decoding across all pairs of linguistic conditions and also across both

scripts for the within-script decoding, and then used a t-test across participants to compare the cross-script and the within-script decoding. We applied repeated measures ANOVA (6 stimuli conditions * 2 groups) to test for pairwise decoding accuracy differences between groups, within each script and each ROI. Effect size was estimated through partial eta-squared (η^2_p) measures. Effect sizes for bilateral LO, V1, and pSTS, can be found in Figures 4-1, 4-3, 5-1, 6-1. Additionally, we applied post-hoc t-tests to assess which conditions drove the effect of conditions (Figures 4-2, 4-4, 5-2, 6-2). Correlations between RDMs were tested through non-parametric statistics. For each correlation between neural matrices, we constructed a null distribution by creating 10000 matrices with shuffled labels and calculating the corresponding correlations. We shuffled both correlation directions and then averaged the distributions for each direction separately and then between directions. The result was a null distribution of 10000 values representing correlations between conditions with shuffled labels. We then compared the correlations obtained from the matrices with intact labels to said distribution and extract a measure of the significance. In the case of correlations between neural-RDMs and our linguistic distance model, we only shuffled labels regarding each subject of one group and the model itself. The final result was again a null distribution of 10000 values representing correlations between conditions with shuffled labels. We then applied the same steps to obtain p-values relative to each observed correlation. The p values in the rmANOVA post-hoc t-test were corrected using Bonferroni correction. All p values in the MVPA analyses were corrected for multiple comparisons with false discovery rate (FDR) corrections.

Results

The Visual Word Form Area of expert readers is selective to visual Braille

The first major question in our study was whether the VWFA as activated by the Latin script is also activated by visual Braille in experts. In all eighteen subjects that took part in the experiment and showed a VWFA for the Latin script (six expert Braille readers, twelve control participants), we performed a General Linear Model (GLM) on the whole-brain level and contrasted univariate activations for intact over scrambled Latin script words (FW > SFW). We then contrasted, in the ROI identified, activity for intact over scrambled Braille words (BW > SBW). In the expert group, individual clusters appeared in the [BW > SBW] contrast, in similar sites to the response to Latin-based stimuli (Figure 2). We performed Small Volume Correction at an individual level by drawing a 10mm radius sphere in the [BW > SBW] contrast around the peak coordinates in VWFA separately and individually identified through the [FW > SFW] contrast. A significant selective response for Braille (BW > SBW) in 5 of the 6 experts emerged in VWFA defined from Latin-based French (Figure 2-1). In the control group, only 1 out of 12 subjects showed a small significant cluster (Figure 3-1). The proportion of activations in participants was significantly different between the two groups ($\text{Chi-square}_{(1)} = 7.03$, $p = 0.008$), showing that people who can read Braille through the visual modality activate VWFA, regardless of the lack of visual features present in canonical scripts (Changizi et al., 2006). Next, we investigated whether a significant preference for intact over scrambled Braille was also present in our expert group as a whole (Figure 2-2), through paired t-tests on the mean activation for intact or

scrambled Braille stimuli. We found a significant difference between intact and scrambled Braille in VWFA defined through the Latin-based script ($t_{(5)}$ = 4.45, p = 0.006). This finding is in line with the aforementioned observation of significant clusters of Braille selectivity in/near Latin-based VWFA. Our results were not influenced by including regressors controlling for eye movements (Figure 2-3) through a decoding of eye motion from fMRI time series using bidsMReye (Gau & Cabee, 2023) and deepMReye (Frey et al., 2021).

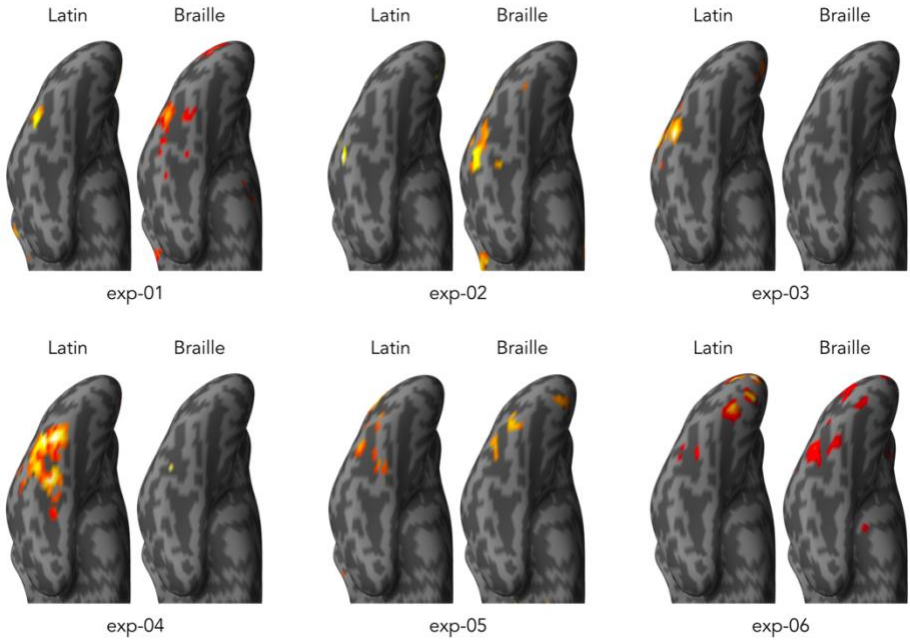

Figure 6 – Visual Word Form Area's univariate response to visual Braille in experts. Results of whole-brain analysis in each expert participant for [FW > SFW] and [BW > SBW] contrasts. For each expert visual Braille reader, contrasts for intact over scrambled Latin stimuli (left side of each participant) and intact over scrambled Braille stimuli (right side of each participant). All results are p_{uncorr} = 0.001 for visualization purposes. Through Small Volume Correction, we identified significant activation for intact over scrambled Braille in sites that localize intact over scrambled Latin stimuli in 5 experts (all except exp-03). Details of the areas localized, with peaks of activity, are available in Figure 2-1. Such selectivity is not present in other areas (Figure 2-2) and is not to be attributed to eye movements (Figure 2-3).

Linguistic properties can be decoded in VWFA when a script is known, regardless of its visual features

To investigate whether the multivariate profiles of VWFA organization was based on linguistic content, we trained a support vector machine (SVM) to classify pairs of linguistic conditions. These conditions (Real Words, Pseudo-Words, Non-Words, Fake Script) differed in the number of linguistic properties presented (see "Methods" section). We decoded pairwise comparisons of linguistic conditions within each script, on an individual level in regions defined by selectivity for intact over scrambled Latin words (Figure 3A). In all the script-group pairs where the stimuli come from a known script (experts seeing Latin-based stimuli, experts seeing Braille stimuli, controls seeing Latin-based stimuli), pairwise decoding accuracy (DA) was on average strongly above chance level (Latin script, expert group: DA = 0.72, $t_{(5)} = 5.850$, $p_{FDR} = 0.002$, D = 2.3; Braille script, expert group: DA = 0.67, $t_{(5)} = 8.386$, $p_{FDR} < 0.001$, D = 3.4; Latin script, control group: DA = 0.72, $t_{(11)} = 5.347$, $p_{FDR} < 0.001$, D = 1.6). In striking contrast, visual Braille decoding in VWFA was not significant in the control group decoding (DA = 0.51, $t_{(11)} = 0.61$, $p_{FDR} = 0.36$). Additionally, we performed two-tailed t-tests on the decoding accuracies (in case of comparisons within a group, we performed paired t-tests) and found significant differences between known and unknown scripts (Latin, experts – Braille, controls: $t_{(8.6)} = 4.73$, $p_{FDR} = 0.002$, D = 2.5; Braille, experts – Braille, controls: $t_{(12.1)} = 5.95$, $p_{FDR} < 0.001$, D = 2.7; Latin, controls – Braille, controls: $t_{(11)} = 4.88$, $p_{FDR} = 0.0012$, D = 1.7, Figure 3A). In contrast, there were no significant differences between conditions across known scripts ($p_{FDR} > 0.36$). These results suggest that the VWFA of experts encodes the differences between linguistic properties in visual Braille as they do, and as controls do, in

their native Latin script, while no such representation is detected for visual Braille in control participants.

Linguistic regularities in VWFA are organized similarly across scripts

Next, we investigated gradients of script-specific representations. We looked in detail at the same pairwise comparisons described in the section above to explore the different relationships between linguistic conditions. We performed within script repeated measures ANOVAs to identify group effects (differences between expert and naïve subjects), effects of the differences in linguistic properties (differences between decoding pairs), and interactions between those two variables (influence of the expertise on the task). In the case of the Braille script, we found an effect of group ($F_{(1)} = 30.825$, $p < 0.001$, $\eta^2_p = 0.66$), as well as an effect of linguistic properties ($F_{(5)} = 6.093$, $p < 0.001$, $\eta^2_p = 0.25$) and an interaction between the two factors ($F_{(5)} = 3.195$, $p = 0.011$, $\eta^2_p = 0.14$), as shown in Figure 3B. These results indicate that the two groups process Braille differently. While VWFA can identify differences between conditions in the case of expert readers, this is not possible in the control group. In contrast, the analysis of the Latin script only showed an effect of the linguistic properties ($F_{(5)} = 11.240$, $p < 0.001$, $\eta^2_p = 0.4$), but no difference between groups ($F_{(1)} = 0.888$, $p = 0.36$) and no interaction ($F_{(5)} = 0.863$, $p = 0.51$). Post-hoc t-tests on the conditions showed that the classification of Real Words against Pseudo-Words (RW-PW) is different (less accurate) from all the other pairs, except the pseudo- and -non-words pair (RW-NW: $p_{Bonf} = 1$, all other $p_{Bonf} < 0.024$). A full table of the post-hoc results is available in Figure 3-2. Differences between conditions show a sensitivity to the linguistic properties of a script such as enhanced linguistic distances triggers enhanced decoding

accuracy for all known scripts (but not for Braille in non-experts). The lack of group differences for the Latin script shows that this sensitivity is induced by differences in expertise in a given script. When groups are compared on their native script (levelling the effect of training), no differences are found. To compare the organizational patterns between scripts and groups, we performed Representational Similarity Analysis (RSA). We built dissimilarity matrices using the pairwise decoding accuracies of each script-group pair (Figure 3C), as well as for a model of linguistic distance that captures the number of linguistic properties that are different between two conditions (with Real Words and Fake Script being most different). We computed the Pearson's correlations between the values in these dissimilarity matrices, and determined their significance through non-parametric permutation analyses (Mattioni et al., 2020; Stelzer et al., 2013) (see "Methods"). We found high correlations for known scripts, both between patterns of neural activity (Latin, experts – Latin, controls: $r = 0.58$, $p_{FDR} < 0.001$; Braille, experts – Latin, controls: $r = 0.58$ $p_{FDR} < 0.001$) and with the linguistic model (Latin, experts – Model: $r = 0.43$, $p_{FDR} = 0.005$; Braille, experts – Model: $r = 0.59$, $p_{FDR} < 0.0001$; Latin, controls – Model: $r = 0.45$, $p_{FDR} < 0.0001$). Correlations with the dissimilarity matrix for Braille in the control group were mostly absent (all $p_{FDR} > 0.20$), with the exception of a correlation with the matrix for Braille stimuli in expert readers ($r = 0.33$, $p_{FDR} = 0.002$). Overall, these correlations further support the view that Braille script in experts is represented in the VWFA in a similar way as Latin-based stimuli and that the type of representation is related to a theoretical model of overlapping linguistic properties inspired by previous findings (Vinckier et al., 2007).

Linguistic representations in VWFA are not generalized across scripts

Since linguistic conditions in Braille readers could be decoded in both scripts (Braille, Latin) we looked at possible generalizations between scripts. We performed cross-script decoding to see if the SVM classifier trained to distinguish linguistic conditions in one script would be able to decode similar conditions in the other script (Figure 3D). A significant result would indicate that the neural patterns associated with the different conditions partially abstract from the visual features intrinsic to each script and would represent linguistic properties independently of the script. We limited this analysis to the expert group since it was the only group showing significant decoding in each script separately. The average cross-script decoding accuracy was not significantly higher than chance (DA = 0.54, $t_{(5)} = 0.58$, $p_{FDR} = 0.40$). This result was maintained in the comparisons of all pairs of linguistic conditions, where no pair of stimuli showed significant cross-script decoding (highest DA value: RW – PW, DA = 0.56, $t_{(5)} = 1.62$, $p_{FDR} = 0.41$). Moreover, we compared the average decoding accuracy within scripts to the cross-decoding accuracy. We found a significant difference ($t_{(5)} = 3.15$, $p = 0.025$), providing positive evidence that the within-script decoding relies upon neural representations that are separated to such a degree that cross-script decoding is affected strongly.

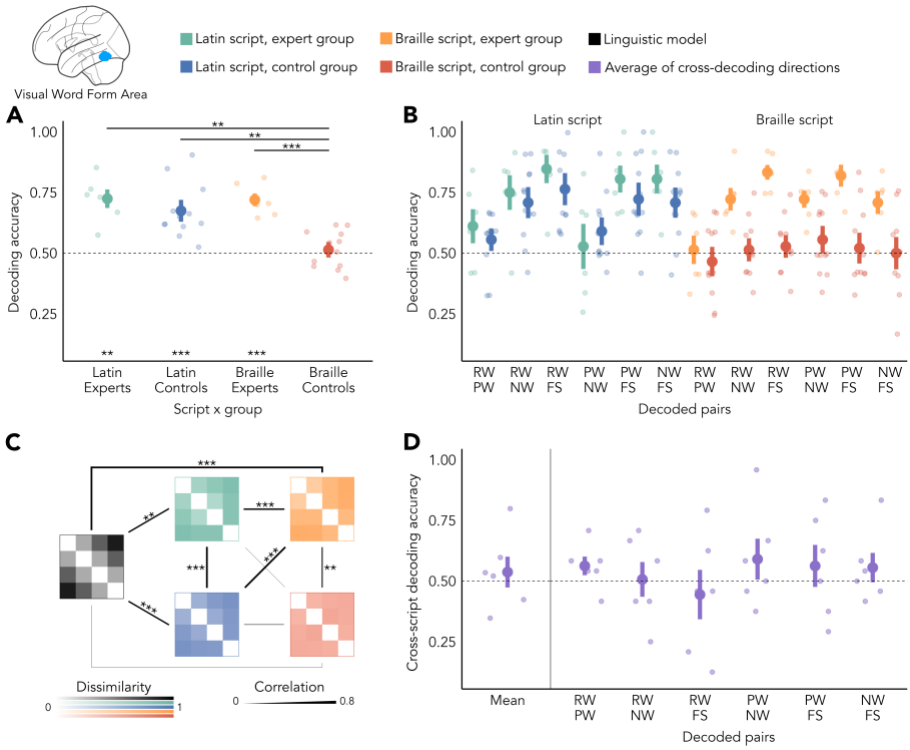

Figure 7 – Multivariate representation of linguistic properties in the Visual Word Form Area (VWFA). (A) Average of pairwise decoding of conditions that differ in linguistic properties. In all the known scripts (Latin and Braille in experts, Latin in controls), we found significantly above-chance decoding and significant differences with the unknown and non-decodable condition of Braille presented to a control group. **(B)** Pairwise decoding within scripts. While experts and controls showed no group difference for their common native script, they showed a group difference for Braille, with experts showing a pattern similar to the one of Latin script. Details of post-hoc analyses can be found in Figure 3-2. **(C)** Representational similarity analysis (RSA) between neural representations (green, orange, blue, red) and with a linguistic model (black). Strength of the correlation is represented by the thickness of the bar connecting two patterns. Organizational patterns of known scripts correlated strongly with each other and with a theoretical model of linguistic distance. A script unknown to the reader (Braille for the control group) does not correlate with the model nor with patterns for Latin script. A significant correlation is present between the neural patterns of Braille. **(D)** Cross-script decoding. The mean of cross-decoding across all pairs of linguistic conditions did not show significant above-chance decoding, and neither did the individual pairs. RW = Real Words; PW = Pseudowords; NW = Nonwords; FS = Fake Script. In all the figures, stars represent significance (* $p \leq 0.05$, ** $p \leq 0.01$, *** $p \leq 0.001$). Error bars represent standard error. Region of interest comes from an example subject. Each ROI was individually localized from peak activations at given contrasts (Table 3-1).

Shape-sensitive Lateral Occipital (LO) does not preferentially process Braille, but represents the differences in linguistic properties

We contrasted fMRI activation for intact and scrambled line drawings to identify individual bilateral Lateral Occipital (LO) sites as two regions of interest (ROIs), one for each hemisphere. To probe for extended plasticity effects of visual Braille in other areas commonly associated with shape, we then performed paired t-tests on the mean activation for intact or scrambled Braille stimuli in those ROIs and found no overall difference in LO (left-LO: $t_{(5)} = 1.50$, $p = 0.19$; right-LO: $t_{(5)} = 1.86$, $p = 0.12$). This result was not dependent on eye movements (Figure 2-3). We then performed the same Multivariate Pattern Analyses applied to VWFA (see "Methods"). In summary, both hemispheres showed similar results as to VWFA. We found significant decoding of linguistic properties (Figs. 4A and 4B) only in scripts known by the groups of participants (left-LO: all $p_{FDR} < 0.006$; right-LO: all $p_{FDR} < 0.011$) and not for Braille processed by the control group (all $p_{FDR} > 0.10$). We also observed statistical differences between known scripts and Braille for controls (all $p_{FDR} < 0.031$). In left-LO, we additionally observed differences between Braille in experts and the native script (both $p_{FDR} < 0.033$). Looking within each script, we found effects in line with VWFA. In Braille, we found group differences (both $p \leq 0.002$), but no effects of linguistic properties (both $p \leq 0.11$) nor interactions (both $p > 0.23$). Within Latin, both hemispheres presented a main effect of linguistic properties (both $p < 0.001$), but no group differences (both $p > 0.36$). In left-LO, we found an interaction between groups and linguistic properties (left-LO: $F_{(5)} = 2.959$, $p = 0.018$; right-LO $F_{(5)} = 0.895$, $p = 0.49$). Detailed statistical tests and results are available in Figure 4-1 and 4-3. These results indicate that experts' LO represents Braille stimuli, although differently from the way Latin-based

stimuli are processed. Representational Similarity Analysis (RSA) on the patterns of pairwise decoding accuracies (Figure 4C) showed strong correlations for known scripts, both between the neural RDMs (left-LO: all $r > 0.55$, all $p_{FDR} < 0.001$. Right-LO: all $r > 0.64$, all $p_{FDR} < 0.001$) and with the linguistic model (all $p_{FDR} \leq 0.002$). In the left-LO of experts, the correlation between Braille-experts RDM with the model was close to significance ($r = 0.34$, $p_{FDR} = 0.052$). Braille representations in controls did not correlate with other neural RDMs or with the model (all $p_{FDR} > 0.07$). No significant cross-script decoding was observed (left-LO: DA = 0.59, $t_{(5)} = 1.56$, $p_{FDR} = 0.16$; right-LO DA = 0.58, $t_{(5)} = 2.35$, $p_{FDR} = 0.10$). Investigating the differences between within scripts and cross-script decoding, we found a significant difference ($t_{(5)} = 3.45$, $p = 0.018$) in the case of right-LO, and an almost-significant trend in the case of left-LO ($t_{(5)} = 2.54$, $p = 0.051$). Finer comparisons were mostly consistent with a lack of cross-decoding and with what we found in VWFA: except for Real Words contrasted to Pseudo-Words in the left LO (DA = 0.56, $t_{(5)} = 4.39$, $p_{FDR} = 0.025$), we observed no significant cross-decoding being above chance level in each pair-wise instances tested.

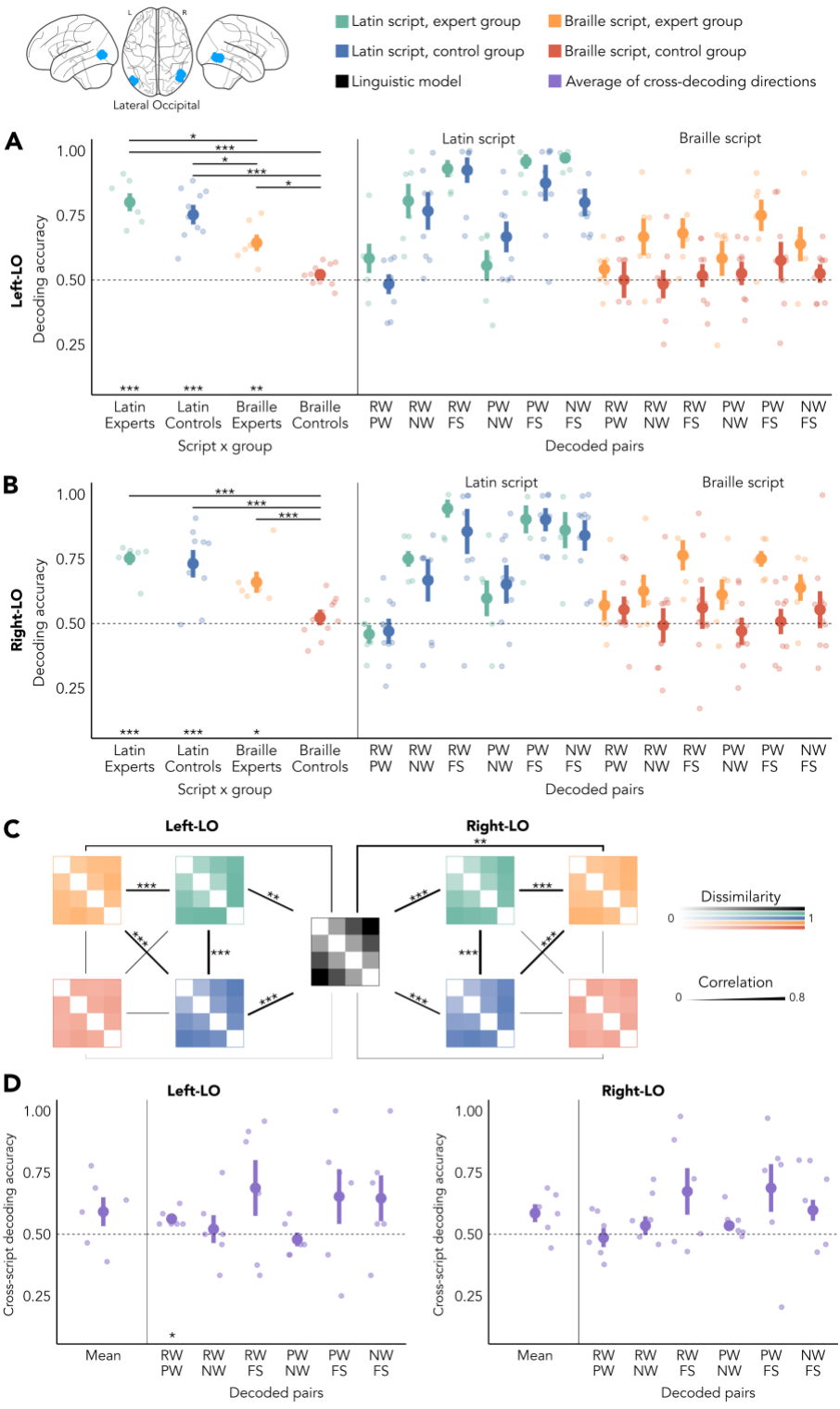
Lateral Occipital
Latin script, expert group
Latin script, control group
Linguistic model
Braille script, expert group
Braille script, control group
Average of cross-decoding directions
A
Left-LO
Decoding accuracy
Latin script
Braille script
Latin Experts
Latin Controls
Braille Experts
Braille Controls
Script x group
Decoded pairs
B
Right-LO
C
Left-LO
Right-LO
Dissimilarity
Correlation
D
Cross-script decoding accuracy
Mean
Decoded pairs

Figure 4 – Multivariate representation of linguistic properties in left and right Lateral Occipital area (LO). **(A)** Pairwise decoding of conditions that differ in linguistic properties in left-LO, average (left) and individual pairs within script (right). Scripts known by the participants showed above-chance decoding accuracy and significant differences with the Braille-experts pair. Significant differences also emerged between the Braille-experts pair and Latin script in both groups. Individual pairs highlighted similar trends within Latin script but different ones within Braille, with experts showing a pattern similar to the one for Latin. **(B)** Pairwise decoding of conditions that differ in linguistic properties in right-LO, average (left) and individual pairs within script (right). Known scripts by the participants showed above-chance decoding accuracy and significant differences with the Braille-experts pair. Individual pairs highlight similar trends within Latin script but different ones within Braille, with experts showing a pattern similar to the one for Latin. Details of post-hoc analyses can be found in Figures 4-2 and 4-4. **(C)** Representational similarity analysis (RSA) between neural representations (green, orange, blue, red) and with a linguistic model (black). Strength of the correlation is represented by the thickness of the bar connecting two patterns. In left-LO, patterns of known scripts correlate strongly with each other. While Latin script patterns also correlate with the model, Braille representations in experts show an almost-significant trend ($p = 0.052$). In right-LO, only neural RDMs relative to known scripts show correlations between themselves and with the linguistic model. **(D)** Cross-script decoding. The mean of cross-decoding across all pairs of linguistic conditions did not show significant above-chance decoding. In right-LO (right plot), no individual comparison, nor the mean, can be decoded across scripts. RW = Real Words; PW = Pseudowords; NW = Nonwords; FS = Fake Script. In all the figures, stars represent significance (* $p \leq 0.05$, ** $p \leq 0.01$, *** $p \leq 0.001$). Error bars represent standard error. Region of interest comes from an example subject. The full list of statistical test for left-LO and right-LO can be found respectively in Figures 4-1 and 4-3.

Multivariate coding of linguistic properties of visual Braille extends to early visual cortex

We further extended our analysis of the organizational pattern found in VWFA and LO, to see if widespread visual Braille processing involves earlier stages of visual perception. To identify the region of interest, we relied on the bilateral early visual cortex (V1) mask from Anatomy toolbox (Eickhoff et al., 2005) and overlapped it with an independent contrast to determine overall visual activity as extracted from the General Linear Model (GLM) for each participant (intact and scrambled Latin words over rest, see “Methods”). Paired t-tests on the mean activation for intact or scrambled Braille stimuli showed no overall difference ($t_{(5)} = 3.67$, $p = 0.07$). This result was confirmed also when we

controlled for eye movements (Figure S2). On the multivariate level (Figure 5A), we found above chance decoding for all scripts (all p_{FDR} <= 0.02). Moreover, decoding accuracies of known scripts differed from participants seeing an unknown script (all p_{FDR} < 0.042). Then, we performed repeated measures ANOVAs for individual scripts. Braille script (Figure 5B) showed a significant main effect of group ($F_{(1)}$ = 10.611, p = 0.005), and of linguistic differences ($F_{(5)}$ = 5.818, p < 0.001), as well as an interaction etween the two ($F_{(5)}$ = 3.011, p = 0.015). Only experts showed distinctions in the neural patterns associated with different linguistic properties of the stimuli. Expectedly, Latin script only showed a main effect of the linguistic properties compared ($F_{(5)}$ = 7.475, p < 0.001), with no differences between groups ($F_{(1)}$ = 0.007, p = 0.94). Post-hoc tests showed that the difference between Real and Pseudo-Words, the classification of stimuli that share the most linguistic features, is significantly different from (lower than) all the other comparisons of conditions (all p_{Bonf} < 0.004). These results indicate that, while Latin script is processed similarly in the two groups, a group difference is present in Braille, surely due to the difference of Braille expertise between our participants. Neural pattern of linguistic distances (Figure 5C) showed significant correlations between the neural RDMs for known scripts (all p_{FDR} < 0.001) and between scripts in the control group (Latin, Controls – Braille, Controls, r = 0.22, p_{FDR} = 0.021), suggesting that the processing of Braille in the expert group is similarly organized to the processing of the Latin script in both groups. The linguistic model correlated significantly with representations of Latin script stimuli (all p_{FDR} < 0.026) but, in contrast to the findings in all other ROIs, not with representations for Braille stimuli (all p_{FDR} > 0.09). Lastly, we did not find any significant cross-script decoding in V1 (Figure 5D), neither in the average

of the comparison pairs (DA: 0.49, $t_{(5)} = -0.36$, $p_{FDR} = 0.86$) nor in the single pairs (most significant value: RW – FS, DA = 0.53, $t_{(5)} = 0.72$, $p_{FDR} = 0.71$), hinting that the distinct low-level visual features of the two scripts result in separate representations. Detailed statistical tests and results are available in Figure 5-1. Finally, there was a significant difference between within-script decoding and cross-script decoding ($t_{(5)} = 3.28$, $p = 0.022$).

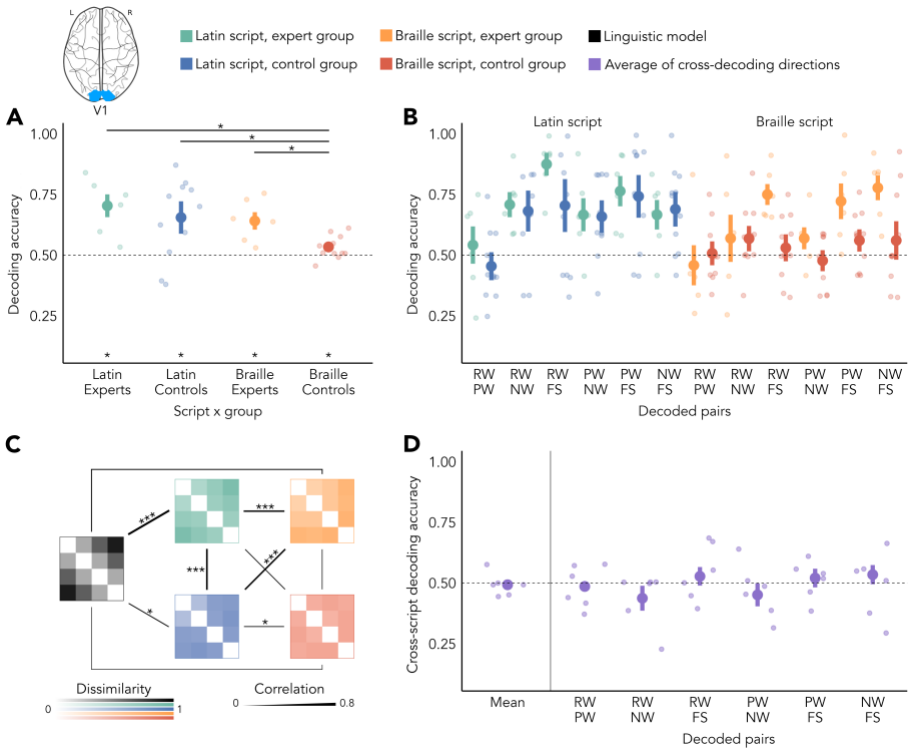

Figure 5 – Multivariate representation of linguistic properties in V1. **(A)** Average of pairwise decoding of conditions that differ in linguistic properties. Stars above the labels on the x-axis represent the statistical significance of the decoding accuracy, while stars above the horizontal bars on top of the graph represent the significance of the difference between conditions (* $p \leq 0.05$, ** $p \leq 0.01$, *** $p \leq 0.001$). Error bars represent standard error. In all the script-group pairs, we found significantly above-chance decoding and significant differences with the unknown and non-decodable condition of Braille presented to a control group. **(B)** Pairwise decoding within scripts. While experts and controls showed no group difference for their common native script, they showed a group difference for Braille, together with a main effect of condition and

interaction between the variables, with experts showing a pattern similar to the one of Latin script. Details of post-hoc analyses can be found in Figure 5-2. **(C)** Representational similarity analysis (RSA) between neural representations (green, orange, blue, red) and with a linguistic model (black). Strength of the correlation is represented by the thickness of the bar connecting two patterns. In V1, patterns of known scripts correlate strongly with each other. Patterns relative to control participants showed a significant correlation across scripts. Only the neural RDMs relative to the organization of the Latin script show a correlation with the linguistic distance model, contrary to what happens in the case of Braille. **(D)** Cross-script decoding. The mean of cross-decoding across all pairs of linguistic conditions did not show significant above-chance decoding, and neither did the individual pairs. RW = Real Words; PW = Pseudowords; NW = Nonwords; FS = Fake Script. Region of interest comes from an example subject. The full list of statistical test can be found in Figure 5-1.

Script-independent representation of linguistic properties in the Left Posterior Temporal Region

Lastly, we observed reliable activations outside of our ROIs when we contrasted intact Latin words over scrambled ones. By masking these activations to regions of the language network (Fedorenko et al., 2010), we identified the left-posterior temporal region, as a higher-order linguistic area. We limited our analyses to those subjects who show univariate activation for Latin words (fifteen participants, five experts). Paired t-test on the mean activation for intact Braille and scrambled dots did not highlight overall preferential activations by intact Braille in the expert group as a whole ($t_{(4)}$ = 2.05, p = 0.11), nor in the control group, as expected ($t_{(10)}$ = -1.24, p = 0.24). These results did not change when eye movements were considered in the GLM. Multivariate analyses within each script revealed an overall profile very similar to the one previously seen in other ROIs. Pairwise comparisons (Figure 6A) could be significantly decoded if the script is known by the participants (all $p_{FDR} < 0.001$), while the script that is unknown to the group had chance-level accuracy (DA = 0.52, $t_{(9)}$ = 0.925, p_{FDR} = 0.24), and significant differences emerged between scripts known and not known by the group of participants

(all $p_{FDR} < 0.001$). More in-depth, pairs of comparison (Figure 6B) showed large differences within Braille (group effect: $F_{(1)} = 58.416$, $p < 0.001$; condition effect: $F_{(5)} = 4.229$, $p = 0.002$), while in the Latin script we observed only an effect of conditions ($F_{(5)} = 4.991$, $p = 0.001$) and no group differences ($F_{(1)} = 0.833$, $p = 0.38$). Post-hoc analyses on the differences between the conditions show that shared linguistic properties are harder to decode (RW-PW and PW-NW, $p_{Bonf} = 0.38$) than more linguistically distant comparisons ($p_{Bonf} < 0.017$). The similarity structure of pairwise comparisons (Figure 6C) showed significant correlations between known scripts (all $p_{FDR} < 0.001$) and no correlation with unknown scripts. Intriguingly, no neural RDM correlated with our linguistic model (all $p_{FDR} > 0.06$). The lack of correlation to the theoretical model suggests different coding principles compared to the other ROIs. Even more striking, and in contrast to all other ROIs, we found the average pairwise cross-decoding accuracy (Figure 6D) to be significantly higher than chance (DA = 0.66, $t_{(5)} = 4.08$, $p_{FDR} = 0.02$, D = 1.8). More in detail, several of the cross-scripts binary decoding were significant (RW-NW, : DA = 0.78, $t_{(5)} = 4.49$, $p_{FDR} = 0.02$; PW-NW: DA = 0.68, $t_{(5)} = 3.64$, $p_{FDR} = 0.025$; PW-FS: DA = 0.68, $t_{(5)} = 2.69$, $p_{FDR} = 0.038$; NW-FS: DA = 0.75, $t_{(5)} = 3.35$, $p_{FDR} = 0.025$). Real Words compared to the Fake Script did not seem to generalize (DA = 0.52, $t_{(5)} = 0.293$, $p_{FDR} = 0.39$). A borderline-significant trend emerged instead in the generalization of the comparison between Real Words and Pseudo-Words (DA = 0.58, $t_{(5)} = 2.36$, $p_{FDR} = 0.052$, D = 1). Together with all this evidence for cross-script decoding, we also observed significant differences with the average within-script decoding accuracy ($t_{(5)} = 3.23$, $p = 0.031$). These findings reveal, overall, a partially script-independent representation of linguistic content. Detailed statistical tests and results are available in Figure 6-1.

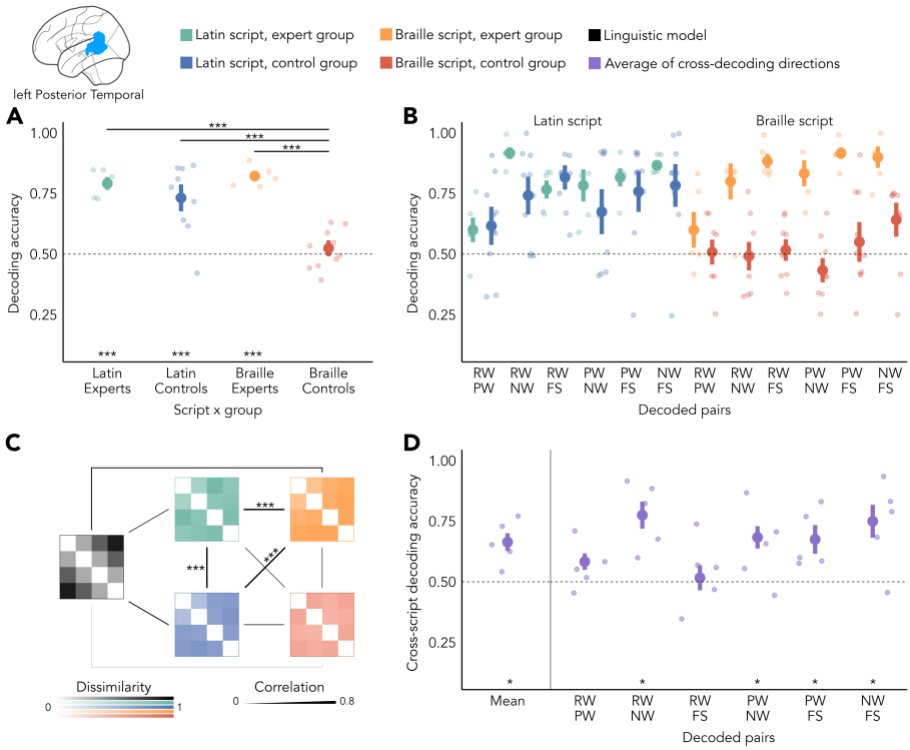

Figure 8 – Multivariate representation of linguistic properties in l-PosTemp. **(A)** Average of pairwise decoding of conditions that differ in linguistic properties. In all the known scripts (Latin and Braille in experts, Latin in controls), we found significantly above-chance decoding and strong significant differences with the unknown and non-decodable condition of Braille presented to a control group. **(B)** Pairwise decoding within scripts. While experts and controls showed no group difference for their common native script, they showed a group difference for Braille, with experts showing a pattern similar to the one of Latin script. Details of post-hoc analyses can be found in Figure 6-2. **(C)** Representational similarity analysis (RSA) between neural representations (green, orange, blue, red) and with a linguistic model (black). Strength of correlation is represented by the thickness of the bar connecting two patterns. Known scripts correlated highly with each other and showed no correlation with the Braille-control pair. No correlation between neural RDMs and the model emerges. **(D)** Cross-script decoding. The mean of cross-decoding across all pairs of linguistic conditions showed significant above-chance decoding, hinting at generalization effects. Different individual pairs can be decoded from one script to another, while RW-PW shows an almost-significant trend ($p = 0.052$). In all the figures, stars represent statistical significance (* $p \leq 0.05$, ** $p \leq 0.01$, *** $p \leq 0.001$). Error bars represent standard error. Region of interest comes from an example subject. The full list of statistical test can be found in Figure 6-1.

Discussion

We investigated the neural bases of processing a peculiar script, visual Braille, that differs from most common scripts for its lack of shape features (Changizi et al., 2006). We found in expert readers a preference for words written in Braille over scrambled arrangements of dots in the visual word form area, as well as multivariate neural representations of linguistic content similar to the one of the participants' native script. Despite these similarities between scripts in the nature of the representations, we did not find generalization of the neural representation between scripts. We also found similar multivariate processing of written text beyond VWFA, in the early stages of visual processing (V1), in shape-sensitive areas (LO), and in higher-order linguistic regions (l-PosTemp). l-PosTemp is the only region showing successful generalization of the two scripts and thus a partially script-invariant representation of linguistic properties. In the visual areas, conversely,

At the univariate level, we found that in expert visual Braille readers, the same site in the ventral occipito-temporal cortex that selectively responds to words presented in the participant's native Latin script also exhibits a preference for Braille words over scrambled dot arrangements. This responsiveness was absent in the control group, suggesting that it can be attributed to neural adaptations in the experts' brains resulting from their training and experience in reading visual Braille. Previous studies showed increased selectivity in VWFA emerging with the acquisition of reading abilities (Dehaene et al., 2015; Dehaene-Lambertz et al., 2018). This sensitivity was investigated in the acquisition of specific scripts (e.g. Hebrew; Baker et al., 2007) and scripts made

out of houses and faces (L. Martin et al., 2019; Moore et al., 2014), suggesting that VWFA could globally link symbols with their represented linguistic properties acquired through learning. Other studies emphasized the importance of shape features for the location and activation of VWFA (Dehaene et al., 2005). However, all investigations were aimed at scripts that are composed of canonical line junctions, and this property of the stimuli may explain why they are co-opted by the reading network. Therefore, none of these studies elucidated whether shape (global form extracted from joined lines) is a mandatory feature to activate VWFA (Ben-Shachar et al., 2007; Roberts et al., 2013; Vogel et al., 2012; Xue & Poldrack, 2007). Other studies investigated Braille, as a unique orthographic system not composed of such explicit shape cues. Previous neuroimaging studies tested tactile Braille, in blind (Büchel et al., 1998; Reich et al., 2011) or sighted (Siuda-Krzywicka et al., 2016) individuals. In blind people, it cannot be excluded that the recruitment of typical orthographic regions can be due to neuroplasticity that occurs in the absence of vision (Xu et al., 2023). In the sighted, activations found for visual Braille could be attributed to mental imagery of classical orthographic letters mapped from tactile Braille (Zhang et al., 2004), a phenomenon arguably less likely when a visual stimulus is presented to the participant. Nevertheless, as a whole these studies already suggested that the domain-specific function of stimuli to read, be it regular scripts, specially trained houses/faces, or even nonvisual scripts, results in a selective response in VWFA for these stimuli. Our findings with visual Braille are consistent with this role of domain-specific task-related factors, and extend them in the direction of a very simple script that lacks the explicit shape cues that are often invoked to explain the location of the VWFA in the visual system.

The similarities between scripts go beyond univariate selectivity and also include the coding principles used to represent the scripts as investigated through multivariate analyses. Differences between patterns of different classes of linguistic stimuli (here Words, Pseudo-Words, Non Words and Fake Scripts) emerged based on the expertise of a script rather than its visual features. In other words, when the visual Braille is learned, it is represented similarly to the native, line-based alphabet. When we consider the patterns of brain activity elicited by our various linguistic conditions, a clear and graded multivoxel selectivity appears evident for visual Braille for conditions with a different lexical status and orthographic properties: Real Words, Pseudo-Words, Non-Words (consonant strings), and scrambled dot patterns. This selectivity was virtually absent in naive participants reading Braille. Furthermore, the coding principles for Braille appeared strikingly similar and strongly correlated to those relative to the Latin-based script.

However, the emerging selectivity for visual Braille is not all about domain specificity and linguistic properties. Our study provided a unique opportunity to investigate how this role of domain specificity interacts with the coding of visual features. Previous studies that showed activity in VWFA for special scripts did not investigate the multivariate organization of these scripts: whether they share common neural representations (suggesting amodal or cross-modal representations) or they co-exist in segregated spaces within the same regions. In our study, we were able to directly compare, in the same participant, orthographic systems with different visual characteristics but common linguistic content. Whereas similar coding principles characterize

visual Braille and Latin alphabet processing, the significant drop in cross-script decoding relative to within-script decoding combined with the absence of cross-scripts generalization (present instead in l-PosTemp), demonstrates that the representations of linguistic properties in VWFA are interdigitated across both scripts. Different neural populations implement the two scripts. These findings resonate with the results of a recent 7T imaging study on French/Chinese bilinguals. In bilinguals, high similarity was present in activated neural populations between French and English, but less so between French and Chinese (Zhan et al., 2023). However, bilinguals might activate different sounds with the two scripts, and the dissociation between French and Chinese could be due to both orthography and phonology. Our study disentangles visual form from phonology and other linguistic factors. While the Latin-based script and visual Braille are visually highly dissimilar, they both map onto the same French words and the same sounds. The total absence of cross-decoding and its significant drop relative to within-script decoding provide clear evidence for the importance of visual form in determining which neural populations in VWFA are selectively activated. The absence of cross-decoding also constraints the interpretation of findings within each script, more than any of the previous studies investigating similar matters (most notably, Vinckier et al., 2007). Within each script there is coding of properties such as lexical access (Words vs other conditions) and statistical regularities (e.g., Pseudo-Words vs Non Words), but in a script-specific coding rather than an abstract linguistic code.

Given that expertise in visual Braille is rare and thus we anticipated a relatively low number of experts, we performed pilot experiments to select two

powerful experimental designs that provide reliable data despite these lower numbers (e.g. very consistent pattern across conditions in the two groups for Latin script in Figs. 3B, 4A, 4B, and 5B, comparisons that can be considered internal replications), and that are able to provide clear data even at an individual level, including univariate activity (see Fig. 2) and high multivariate decoding accuracies (e.g., very clear split between experts and controls on Braille in Fig. 3A). As a result, the expertise-related effects are striking. However, the null result of finding no cross-decoding could depend on the number of participants. It is possible that a study with a much larger group of experts would in the end find a cross-decoding that is still small but significant. Here it is important that the conclusions about independent, script-specific representations are also fully supported by the significant drop in cross-decoding relative to within-script decoding. This is a clear effect and not a null result. Furthermore, the findings in l-PosTemp reveal that it is possible to find significant cross-decoding with our design and analyses methods. Together, our findings indicate that script specificity is a major property of the expertise-related representations in VWFA.

Similarly, the choice of an implicit task reveals the strength of the results. To be able to accurately compare the responses of the two groups, we opted for a task (1-back, see Methods) that could be performed by both groups given with the same instruction. It is likely that naive Braille readers engaged in a purely visual task when facing Braille stimuli while the expert participants would additionally employ a linguistic strategy. The presentation of Latin and Braille alphabets in different run, reduces the contextual effects between them.

Other visual areas showed no univariate differences between Braille and Scrambled Braille bilaterally in V1 and shape-selective LO in expert Braille readers. However, in expert readers, these regions showed similar multivariate profiles to the one identified in VWFA. This similarity in the patterns of linguistic distance between both known scripts reveals distributed expertise-induced plasticity in the visual system. Studies of visual word forms focused mainly on VWFA, but more widespread linguistic selectivity have already been observed for shape-based scripts across the occipito-temporal cortices (Szwed et al., 2014; Siuda-Krzywicka et al., 2016). Although V1 has already been associated with plasticity effects beyond VWFA and preference for known scripts (e.g. Latin-based and Chinese; Szwed et al., 2014), it is remarkable that such adaptation extends to a script that lacks explicit shape cues. Similarly, it is striking that LO shows significant coding of word-like stimuli, and that such coding is found in both hemispheres. Previous studies on the gradient of involvement in word processing (Vinckier et al., 2007) did not perform such multivariate analyses, so possibly the literature underestimates the involvement of LO in word processing. Indeed the multivariate profile in LO is not a simple consequence of visual differences across classes of stimuli, as this representation was not observed in control participants for Braille; nor is it restricted to typical script shape processing, as we also observed this coding in visual Braille for experts. This is particularly striking because LO is often seen as the major region supporting shape perception (Grill-Spector et al., 2001; Kanwisher et al., 1996; Malach et al., 1995), and visual Braille letters contain much less shape cues compared to other visual scripts. We note however that the absence of explicit shape cues does not exclude the possibility that the visual system would still engage in forms of pattern recognition, like Gestalt

principles, such as proximity and good continuation, that would include the perception of subjective line and shape elements that would connect the dots with lines and more complex geometric patterns. However these principles might also emerge in the non-expert group, therefore not explaining our results. Also, previous studies have suggested that the Braille script is more difficult to learn than line or shape-based scripts, also suggesting the idiosyncrasy of Braille compared to line and shaped based scripts (Bola et al., 2017).

Beyond visual regions, we reliably identified activated voxels in the left posterior temporal area (l-PosTemp) with patterns of multivariate activity that resemble the ones found in visual areas. Although the neural patterns for known scripts presented strong correlations with each other, we surprisingly found no correlation with the linguistic distance model. The absence of correlations may be attributed to the model, inspired by results shown in vOTC (Vinckier et al., 2007), a region highly sensible to visual input. It is possible that the role of posterior temporal regions in phonological processing (Bhaya-Grossman & Chang, 2022; Chang et al., 2010) and reading (A. Martin et al., 2015) determines non-linear relationships between linguistic properties. The preferential phonological tuning of this region could lead to increased similarity between Words and Pseudo-Words, and increased dissimilarity to low-frequency phonological categories (Non-Words and Fake Script). Finally, the cross-script generalization indicates that the representations of linguistic conditions are invariant to the visual differences. Such common, script-invariant representation further demonstrates that at this stage of information processing in l-PosTemp, linguistic information is abstracted from the visual

properties of the stimuli processed in the ventral occipito-temporal cortex (A. Martin et al., 2015). Note though that this script invariance was only partial, because the cross-script decoding was significantly smaller than the within-script decoding.

Overall, we observe that several regions of the reading network are involved in the processing of the orthographic material in a script without explicit line junctions. Looking beyond the domain of reading, our work adds to the literature showing effects of learning in visual processing, including learning the statistical properties of the stimuli such as manipulated in the comparison of Pseudo-Words and Non Words. Such learning has been found in other domains, namely the recognition of chessboards in professional chess players (Bilalic et al., 2011) or the processing of Greeble stimuli (Gauthier et al., 1998; Gauthier & Tarr, 1997) in the fusiform face area (FFA), proving the ability of VOTC to adapt to an extended set of stimuli. We show that a coding of the statistical regularities that compose the stimuli extends beyond the canonical line-based script and also apply to visual Braille, a script that is void of those explicit shape cues. Such impact does not come as a surprise, as statistical learning is part of the language acquisition process, but it is remarkable that it extends to this broad variety of visual features.

In conclusion, despite marked differences in the visual properties of Braille and Latin-based scripts, the coding principles of occipital regions showing a preference for a Latin-based script are strikingly similar to responses to visual Braille in experts readers. In occipital regions the neuronal populations that implement these principles do not overlap, suggesting interdigitated

representation across scripts co-existing in the same region (Zhan et al., 2023); while in posterior temporal regions scripts are represented in a more script-invariant manner.

Data and code availability

All the code used to perform the experiments (https://github.com/fcerpe/VBE_experiment) and analyse the results (https://github.com/fcerpe/VBE_data) is publicly available on GitHub. Raw defaced MRI scans are available on GIN in BIDS (Gorgolewski et al., 2016) format (https://gin.g-node.org/cpp_brewery/2022_StLuc_VisBraExpertise_FC_raw).

Author contributions

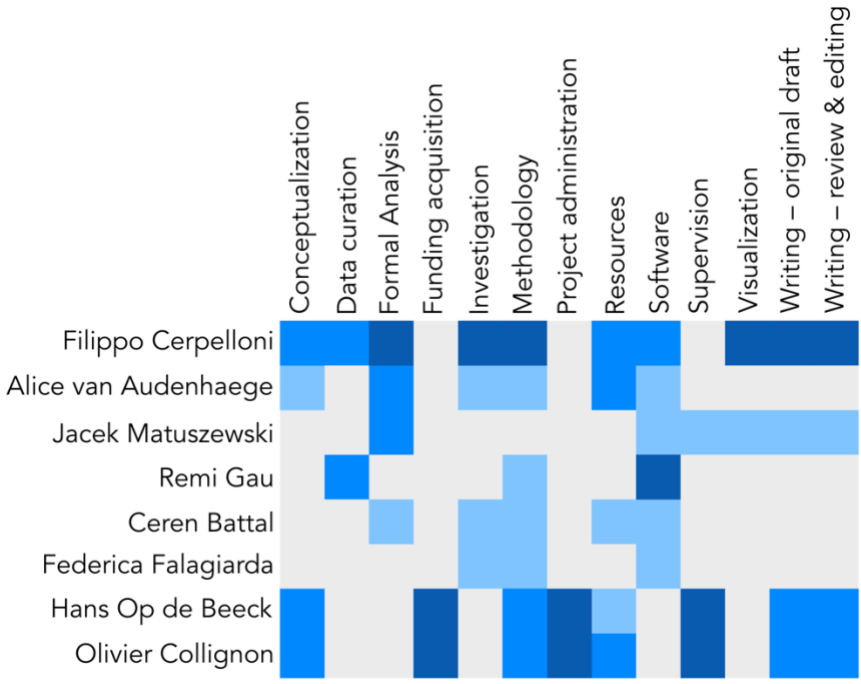

Figure 9 – Author contributions. Based on the CRedIT (*https://credit.niso.org/*) taxonomy. Each contribution was assessed as 'lead' (dark blue), 'equal' (mid blue), 'support' (light blue).

Acknowledgements

The authors would like to thank the Institut Royal pour Sourds et Aveugles (IRSA) and the association EQLA for their help in recruiting the Braille experts. The project was funded in parts by a Mandat d'Impulsion Scientifique awarded to OC, the Belgian Excellence of Science (EOS) program (Project No. 30991544) awarded to HO and OC, a Flagship ERA-NET grant SoundSight (FRS-FNRS PINT-MULTI R.8008.19) awarded to OC, and research project G0D3322N of the Fund for Scientific Research (FWO) Flanders and Methusalem project METH/24/003 awarded to HO. AV is a research fellow, CB and JM are postdoctoral researchers, and OC is a senior research associate at the Fond National de la Recherche

Scientifique de Belgique (FRS-FNRS). FC received funding from a KU Leuven-UCLouvain cooperation grant.

Conflict of interest statement

The authors declare no competing financial interests.

References

Arcaro, M. J., & Livingstone, M. S. (2017). A hierarchical, retinotopic proto-organization of the primate visual system at birth. *eLife*, *6*, e26196. https://doi.org/10.7554/eLife.26196

Baker, C. I., Liu, J., Wald, L. L., Kwong, K. K., Benner, T., & Kanwisher, N. (2007). Visual word processing and experiential origins of functional selectivity in human extrastriate cortex. *Proceedings of the National Academy of Sciences*, *104*(21), 9087–9092. https://doi.org/10.1073/pnas.0703300104

Ben-Shachar, M., Dougherty, R. F., Deutsch, G. K., & Wandell, B. A. (2007). Differential Sensitivity to Words and Shapes in Ventral Occipito-Temporal Cortex. *Cerebral Cortex*, *17*(7), 1604–1611. https://doi.org/10.1093/cercor/bhl071

Bhaya-Grossman, I., & Chang, E. F. (2022). Speech Computations of the Human Superior Temporal Gyrus. *Annual Review of Psychology*, *73*(1), 79–102. https://doi.org/10.1146/annurev-psych-022321-035256

Biederman, I. (1987). *Recognition-by-Components: A Theory of Human Image Understanding*.

Bilalic, M., Langner, R., Ulrich, R., & Grodd, W. (2011). Many Faces of Expertise: Fusiform Face Area in Chess Experts and Novices. *Journal of Neuroscience*, *31*(28), 10206–10214. https://doi.org/10.1523/JNEUROSCI.5727-10.2011

Bola, Ł., Radziun, D., Siuda-Krzywicka, K., Sowa, J. E., Paplińska, M., Sumera, E., & Szwed, M. (2017). Universal Visual Features Might Be Necessary for Fluent Reading. A Longitudinal Study of Visual Reading in Braille and Cyrillic Alphabets. *Frontiers in Psychology*, *8*. https://doi.org/10.3389/fpsyg.2017.00514

Brodeur, M. B., Dionne-Dostie, E., Montreuil, T., & Lepage, M. (2010). The Bank of Standardized Stimuli (BOSS), a New Set of 480 Normative Photos of Objects to Be Used as Visual Stimuli in Cognitive Research. *PLoS ONE*, *5*(5), e10773. https://doi.org/10.1371/journal.pone.0010773

Büchel, C., Price, C., & Friston, K. (1998). A multimodal language region in the ventral visual pathway. *Nature*, *394*(6690), 274–277. https://doi.org/10.1038/28389

Carling, K. (2000). Resistant outlier rules and the non-Gaussian case. *Computational Statistics & Data Analysis*, *33*(3), 249–258. https://doi.org/10.1016/S0167-9473(99)00057-2

Chang, E. F., Rieger, J. W., Johnson, K., Berger, M. S., Barbaro, N. M., & Knight, R. T. (2010). Categorical speech representation in human superior temporal gyrus. *Nature Neuroscience*, *13*(11), 1428–1432. https://doi.org/10.1038/nn.2641

Changizi, M. A., Zhang, Q., Ye, H., & Shimojo, S. (2006). *The Structures of Letters and Symbols throughout Human History Are Selected to Match Those Found in Objects in Natural Scenes.* 23.

Code braille français uniformisé pour la transcription des textes imprimés (CBFU) Réalisé. (2008). *76*(3), 61–64.

Cohen, L., Dehaene, S., Naccache, L., Lehéricy, S., Dehaene-Lambertz, G., Hénaff, M.-A., & Michel, F. (2000). The visual word form area: Spatial and temporal characterization of an initial stage of reading in normal subjects and posterior split-brain patients. *Brain: A Journal of Neurology*, *123*, 291–307.

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

Coquillart, N., Collignon, O., Cerpelloni, F., & Van Audenhaege, A. (2022). *Lecture visuelle du braille et plasticité cérébrale. Zoom sur l'aire visuelle de la forme des mots* [UCLouvain]. https://dial.uclouvain.be/memoire/ucl/object/thesis:37423

Dębska, A., Wójcik, M., Chyl, K., Dzięgiel-Fivet, G., & Jednoróg, K. (2023). Beyond the Visual Word Form Area – a cognitive characterization of the left ventral occipitotemporal cortex. *Frontiers in Human Neuroscience*, *17*, 1199366. https://doi.org/10.3389/fnhum.2023.1199366

Dehaene, S., & Cohen, L. (2011). The unique role of the visual word form area in reading. *Trends in Cognitive Sciences*, *15*(6), 254–262. https://doi.org/10.1016/j.tics.2011.04.003

Dehaene, S., Cohen, L., Morais, J., & Kolinsky, R. (2015). Illiterate to literate: Behavioural and cerebral changes induced by reading acquisition.

Nature Reviews Neuroscience, *16*(4), 234–244. https://doi.org/10.1038/nrn3924

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

Dehaene-Lambertz, G., Monzalvo, K., & Dehaene, S. (2018). The emergence of the visual word form: Longitudinal evolution of category-specific ventral visual areas during reading acquisition. *PLOS Biology*, *16*(3), e2004103. https://doi.org/10.1371/journal.pbio.2004103

Eickhoff, S. B., Stephan, K. E., Mohlberg, H., Grefkes, C., Fink, G. R., Amunts, K., & Zilles, K. (2005). A new SPM toolbox for combining probabilistic cytoarchitectonic maps and functional imaging data. *NeuroImage*, *25*(4), 1325–1335. https://doi.org/10.1016/j.neuroimage.2004.12.034

Fedorenko, E., Behr, M. K., & Kanwisher, N. (2011). Functional specificity for high-level linguistic processing in the human brain. *Proceedings of the National Academy of Sciences*, *108*(39), 16428–16433. https://doi.org/10.1073/pnas.1112937108

Fedorenko, E., Hsieh, P.-J., Nieto-Castañón, A., Whitfield-Gabrieli, S., & Kanwisher, N. (2010). New Method for fMRI Investigations of Language: Defining ROIs Functionally in Individual Subjects. *Journal of Neurophysiology*, *104*(2), 1177–1194. https://doi.org/10.1152/jn.00032.2010

Ferrand, L., New, B., Brysbaert, M., Keuleers, E., Bonin, P., Méot, A., Augustinova, M., & Pallier, C. (2010). The French Lexicon Project: Lexical decision data for 38,840 French words and 38,840

pseudowords. *Behavior Research Methods*, *42*(2), 488–496. https://doi.org/10.3758/BRM.42.2.488

Frey, M., Nau, M., & Doeller, C. F. (2021). Magnetic resonance-based eye tracking using deep neural networks. *Nature Neuroscience*, *24*(12), 1772–1779. https://doi.org/10.1038/s41593-021-00947-w

Gau, R., Barilari, M., Battal, C., Caron-Guyon, J., Cerpelloni, F., Falagiarda, F., MacLean, M., Mattioni, S., Rezk, M., Shahzad, I., Yang, Y., & Collignon, O. (2023). *Bidspm: An spm-centric bids app for flexible statistical analysis*. OHBM. https://dial.uclouvain.be/pr/boreal/en/object/boreal%3A279516

Gau, R., & Cabee, P. (2023). *bidsMReye* (Version 0.3.0) [Computer software]. Zenodo. https://doi.org/10.5281/zenodo.7574501

Gauthier, I., & Tarr, M. J. (1997). Becoming a "Greeble" Expert: Exploring Mechanisms for Face Recognition. *Vision Research*, *37*(12), 1673–1682. https://doi.org/10.1016/S0042-6989(96)00286-6

Gauthier, I., Williams, P., Tarr, M. J., & Tanaka, J. (1998). Training 'greeble' experts: A framework for studying expert object recognition processes. *Vision Research*, *38*(15–16), 2401–2428. https://doi.org/10.1016/S0042-6989(97)00442-2

Gimenes, M., Perret, C., & New, B. (2020). Lexique-Infra: Grapheme-phoneme, phoneme-grapheme regularity, consistency, and other sublexical statistics for 137,717 polysyllabic French words. *Behavior Research Methods*, *52*(6), 2480–2488. https://doi.org/10.3758/s13428-020-01396-2

Gorgolewski, K. J., Auer, T., Calhoun, V. D., Craddock, R. C., Das, S., Duff, E. P., Flandin, G., Ghosh, S. S., Glatard, T., Halchenko, Y. O., Handwerker, D.

A., Hanke, M., Keator, D., Li, X., Michael, Z., Maumet, C., Nichols, B. N., Nichols, T. E., Pellman, J., … Poldrack, R. A. (2016). The brain imaging data structure, a format for organizing and describing outputs of neuroimaging experiments. *Scientific Data*, *3*(1), 160044. https://doi.org/10.1038/sdata.2016.44

Grill-Spector, K., Kourtzi, Z., & Kanwisher, N. (2001). The lateral occipital complex and its role in object recognition. *Vision Research*, *41*(10–11), 1409–1422. https://doi.org/10.1016/S0042-6989(01)00073-6

Grill-Spector, K., & Malach, R. (2004). The human visual cortex. *Annual Review of Neuroscience*, *27*(1), 649–677. https://doi.org/10.1146/annurev.neuro.27.070203.144220

Kanwisher, N., Chun, M. M., McDermott, J., & Ledden, P. J. (1996). Functional imaging of human visual recognition. *Cognitive Brain Research*, *5*(1–2), 55–67. https://doi.org/10.1016/S0926-6410(96)00041-9

Kriegeskorte, N., Mur, M., & Bandettini, P. (2008). Representational similarity analysis—Connecting the branches of systems neuroscience. *Frontiers in Systems Neuroscience*, *2*(NOV), 1–28. https://doi.org/10.3389/neuro.06.004.2008

Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2017). ImageNet classification with deep convolutional neural networks. *Communications of the ACM*, *60*(6), 84–90. https://doi.org/10.1145/3065386

Landau, B., Smith, L. B., & Jones, S. S. (1988). The importance of shape in early lexical learning. *Cognitive Development*, *3*(3), 299–321. https://doi.org/10.1016/0885-2014(88)90014-7

Li, J., Osher, D. E., Hansen, H. A., & Saygin, Z. M. (2020). Innate connectivity patterns drive the development of the visual word form area. *Scientific Reports*, *10*(1), 18039. https://doi.org/10.1038/s41598-020-75015-7

Malach, R., Reppas, J. B., Benson, R. R., Kwong, K. K., Jiang, H., Kennedy, W. A., Ledden, P. J., Brady, T. J., Rosen, B. R., & Tootell, R. B. (1995). Object-related activity revealed by functional magnetic resonance imaging in human occipital cortex. *Proceedings of the National Academy of Sciences*, *92*(18), 8135–8139. https://doi.org/10.1073/pnas.92.18.8135

Marr, D., & Hildreth, E. (1980). *Theory of edge detection*.

Martin, A., Schurz, M., Kronbichler, M., & Richlan, F. (2015). Reading in the brain of children and adults: A meta-analysis of 40 functional magnetic resonance imaging studies. *Human Brain Mapping*, *36*(5), 1963–1981. https://doi.org/10.1002/hbm.22749

Martin, L., Durisko, C., Moore, M. W., Coutanche, M. N., Chen, D., & Fiez, J. A. (2019). The VWFA Is the Home of Orthographic Learning When Houses Are Used as Letters. *Eneuro*, *6*(1), ENEURO.0425-17.2019. https://doi.org/10.1523/ENEURO.0425-17.2019

Mattioni, S., Rezk, M., Battal, C., Bottini, R., Cuculiza Mendoza, K. E., Oosterhof, N. N., & Collignon, O. (2020). Categorical representation from sound and sight in the ventral occipito-temporal cortex of sighted and blind. *eLife*, *9*, e50732. https://doi.org/10.7554/eLife.50732

Moore, M. W., Durisko, C., Perfetti, C. A., & Fiez, J. A. (2014). Learning to Read an Alphabet of Human Faces Produces Left-lateralized Training Effects in the Fusiform Gyrus. *Journal of Cognitive Neuroscience*, *26*(4), 896–913. https://doi.org/10.1162/jocn_a_00506

New, B., Pallier, C., Ferrand, L., & Matos, R. (2005). Manuel de Lexique 3. *Behavior Research Methods, Instruments, & Computers*, *36*(3), 516–524.

Oosterhof, N. N., Connolly, A. C., & Haxby, J. V. (2016). CoSMoMVPA: Multi-Modal Multivariate Pattern Analysis of Neuroimaging Data in Matlab/GNU Octave. *Frontiers in Neuroinformatics*, *10*. https://doi.org/10.3389/fninf.2016.00027

Pattamadilok, C., Planton, S., & Bonnard, M. (2019). Spoken language coding neurons in the Visual Word Form Area: Evidence from a TMS adaptation paradigm. *NeuroImage*, *186*, 278–285. https://doi.org/10.1016/j.neuroimage.2018.11.014

Planton, S., Chanoine, V., Sein, J., Anton, J.-L., Nazarian, B., Pallier, C., & Pattamadilok, C. (2019). Top-down activation of the visuo-orthographic system during spoken sentence processing. *NeuroImage*, *202*, 116135. https://doi.org/10.1016/j.neuroimage.2019.116135

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

Price, C. J., & Devlin, J. T. (2004). The pro and cons of labelling a left occipitotemporal region: "The visual word form area". *NeuroImage*, *22*(1), 477–479. https://doi.org/10.1016/j.neuroimage.2004.01.018

Price, C. J., & Devlin, J. T. (2011). The Interactive Account of ventral occipitotemporal contributions to reading. *Trends in Cognitive Sciences*, *15*(6), 246–253. https://doi.org/10.1016/j.tics.2011.04.001

Rajalingham, R., Kar, K., Sanghavi, S., Dehaene, S., & DiCarlo, J. J. (2020). The inferior temporal cortex is a potential cortical precursor of

orthographic processing in untrained monkeys. *Nature Communications*, *11*(1), 3886. https://doi.org/10.1038/s41467-020-17714-3

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

Riesenhuber, M., & Poggio, T. (1999). Hierarchical models of object recognition in cortex. *Nature Neuroscience*, *2*(11), 1019–1025. https://doi.org/10.1038/14819

Roberts, D. J., Woollams, A. M., Kim, E., Beeson, P. M., Rapcsak, S. Z., & Lambon Ralph, M. A. (2013). Efficient Visual Object and Word Recognition Relies on High Spatial Frequency Coding in the Left Posterior Fusiform Gyrus: Evidence from a Case-Series of Patients with Ventral Occipito-Temporal Cortex Damage. *Cerebral Cortex*, *23*(11), 2568–2580. https://doi.org/10.1093/cercor/bhs224

Rossion, B., & Pourtois, G. (2004). Revisiting Snodgrass and Vanderwart's Object Pictorial Set: The Role of Surface Detail in Basic-Level Object Recognition. *Perception*, *33*(2), 217–236. https://doi.org/10.1068/p5117

Saygin, Z. M., Osher, D. E., Norton, E. S., Youssoufian, D. A., Beach, S. D., Feather, J., Gaab, N., Gabrieli, J. D. E., & Kanwisher, N. (2016). Connectivity precedes function in the development of the visual word form area. *Nature Neuroscience*, *19*(9), 1250–1255. https://doi.org/10.1038/nn.4354

Siuda-Krzywicka, K., Bola, Ł., Paplińska, M., Sumera, E., Jednoróg, K., Marchewka, A., Śliwińska, M. W., Amedi, A., & Szwed, M. (2016).

Massive cortical reorganization in sighted Braille readers. *eLife*, *5*, e10762. https://doi.org/10.7554/eLife.10762

Stelzer, J., Chen, Y., & Turner, R. (2013). Statistical inference and multiple testing correction in classification-based multi-voxel pattern analysis (MVPA): Random permutations and cluster size control. *NeuroImage*, *65*, 69–82. https://doi.org/10.1016/j.neuroimage.2012.09.063

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

Szwed, M., Dehaene, S., Kleinschmidt, A., Eger, E., Valabrègue, R., Amadon, A., & Cohen, L. (2011). Specialization for written words over objects in the visual cortex. *NeuroImage*, *56*(1), 330–344. https://doi.org/10.1016/j.neuroimage.2011.01.073

Szwed, M., Qiao, E., Jobert, A., Dehaene, S., & Cohen, L. (2014). Effects of Literacy in Early Visual and Occipitotemporal Areas of Chinese and French Readers. *Journal of Cognitive Neuroscience*, *26*(3), 459–475. https://doi.org/10.1162/jocn_a_00499

Tarr, M. J., & Bulthoff, H. H. (1998). *Image-based object recognition in man, monkey and machine*.

Vinckier, F., Dehaene, S., Jobert, A., Dubus, J. P., Sigman, M., & Cohen, L. (2007). Hierarchical Coding of Letter Strings in the Ventral Stream: Dissecting the Inner Organization of the Visual Word-Form System. *Neuron*, *55*(1), 143–156. https://doi.org/10.1016/j.neuron.2007.05.031

Vogel, A. C., Petersen, S. E., & Schlaggar, B. L. (2012). The Left Occipitotemporal Cortex Does Not Show Preferential Activity for Words. *Cerebral Cortex*, *22*(12), 2715–2732. https://doi.org/10.1093/cercor/bhr295

Xu, Y., Vignali, L., Sigismondi, F., Crepaldi, D., Bottini, R., & Collignon, O. (2023). Similar object shape representation encoded in the inferolateral occipitotemporal cortex of sighted and early blind people. *PLOS Biology*, *21*(7), e3001930. https://doi.org/10.1371/journal.pbio.3001930

Xue, G., & Poldrack, R. A. (2007). The Neural Substrates of Visual Perceptual Learning of Words: Implications for the Visual Word Form Area Hypothesis. *Journal of Cognitive Neuroscience*, *19*(10), 1643–1655. https://doi.org/10.1162/jocn.2007.19.10.1643

Yarkoni, T., Poldrack, R. A., Nichols, T. E., Van Essen, D. C., & Wager, T. D. (2011). Large-scale automated synthesis of human functional neuroimaging data. *Nature Methods*, *8*(8), 665–670. https://doi.org/10.1038/nmeth.1635

Zhan, M., Pallier, C., Agrawal, A., Dehaene, S., & Cohen, L. (2023). Does the visual word form area split in bilingual readers? A millimeter-scale 7-T fMRI study. *Science Advances*.

Zhang, M., Weisser, V. D., Stilla, R., Prather, S. C., & Sathian, K. (2004). Multisensory cortical processing of object shape and its relation to mental imagery. *Cognitive, Affective, & Behavioral Neuroscience*, *4*(2), 251–259. https://doi.org/10.3758/CABN.4.2.251

Appendix

Table 2-1 – Functional activation of Visual Word Form Area (VWFA) in expert visual Braille readers

Contrast	Subject	p-value (FWE corrected)	z-score	Cluster size	MNI coordinates
FW > SFW	exp-01	0.000	5.04	24	-47 -60 -13
	exp-02	0.000	4.78	9	-52 -47 -18
	exp-03	0.000	6.57	83	-52 -55 -18
	exp-04	0.000	5.90	159	-44 -63 -13
	exp-05	0.002	4.31	29	-44 -60 -15
	exp-06	0.000	5.27	27	-44 -60 -13
BW > SBW	exp-01	0.000	5.09	113	-49 -63 -13
	exp-02	0.000	5.23	92	-55 -47 -21
	exp-03	NA	NA	NA	NA
	exp-04	0.034	3.50	7	-39 -60 -5
	exp-05	0.003	4.23	37	-42 -57 -8
	exp-06	0.000	5.39	79	-47 -63 -15

Table 2-1 – Functional activation of Visual Word Form Area (VWFA) in expert visual Braille readers. For each expert, we contrasted intact Latin words (FW) to scrambled ones (SFW) and obtained activation peaks. To determine the cluster size of VWFA, we performed Small Volume Correction in a 10mm radius sphere around the individual peak coordinates of VWFA identified through a Latin-script contrast (FW > SFW). Next, we performed an additional Small Volume Correction around the same peaks to determine significantly activated voxels for the contrast of Braille words (BW) over scrambled dots (SBW). MNI coordinates refer to the peak of each contrast.

Figure 2-2 – Univariate selectivity for intact over scrambled Braille across ROIs

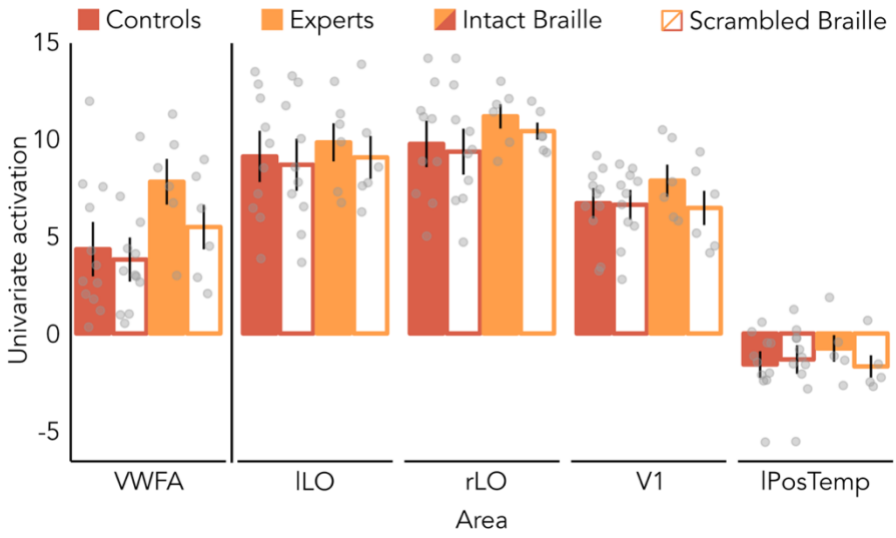

Figure 2-2 - Univariate selectivity for intact over scrambled Braille across ROIs. For each ROI, we extracted the mean activation individually for intact Braille (full bars) and scrambled Braille (empty bars) stimuli. Black bars represent standard error (SE), grey dots represent individual subjects.

Figure 2-3 - Univariate selectivity for intact over scrambled Braille across ROIs with eye movements

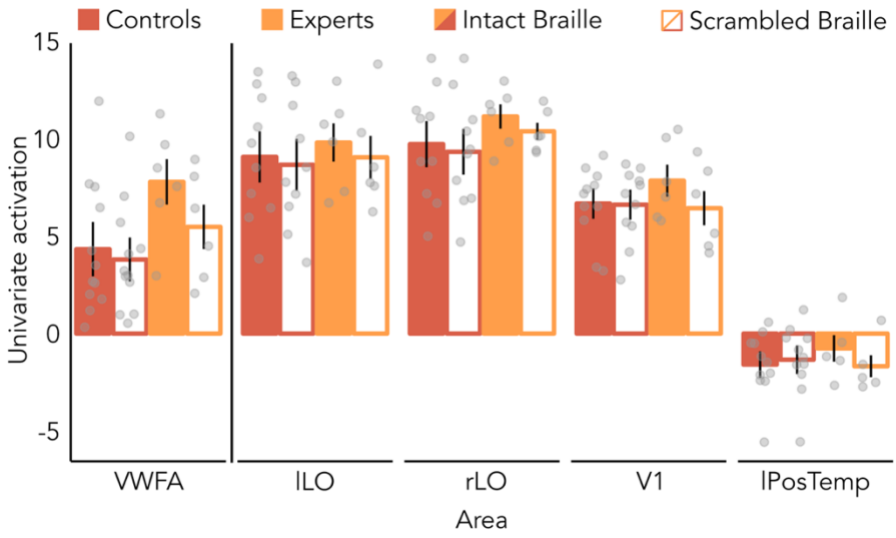

Figure 2-3 - Univariate selectivity for intact over scrambled Braille across ROIs with eye movements. For each ROI, we extracted the mean activation individually for intact Braille (full bars) and scrambled Braille (empty bars) stimuli, using the GLM analysis where eye movements (computed through bidsMReye toolbox). Black bars represent standard error (SE), grey dots represent individual subjects.

Figure 3-1 – Comparison of ROI definitions and their multivariate representation of linguistic properties in the Visual Word Form Area (VWFA)

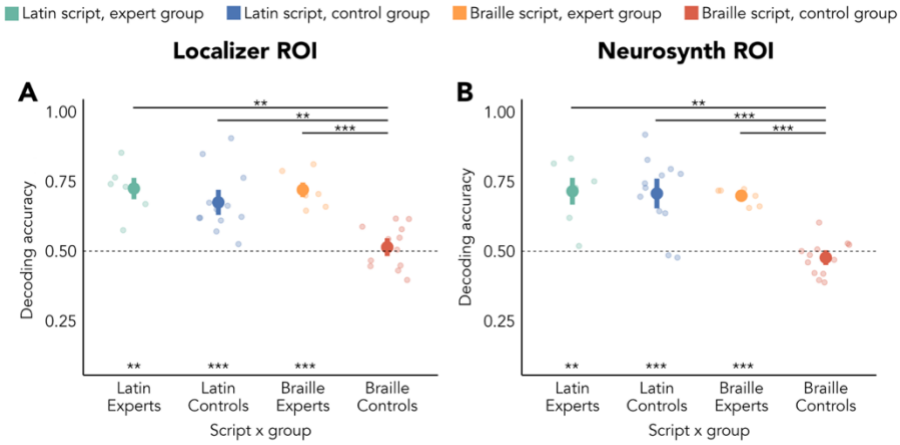

Figure 10-1 – Comparison of ROI definitions and their multivariate representation of linguistic properties in the Visual Word Form Area (VWFA). (A) Region of Interest defined through the methods reported in "Definition of the Regions of Interest (ROIs)". (B) Region of Interest defined through Neurosynth (keyword: "visual words"). We selected the highest-responding 100 voxels and applied the mask to each subject, using the same analysis pipeline used for our custom definition.

Figure 3-2 – Correlations between neural and model RDMs in the Visual Word Form Area (VWFA)

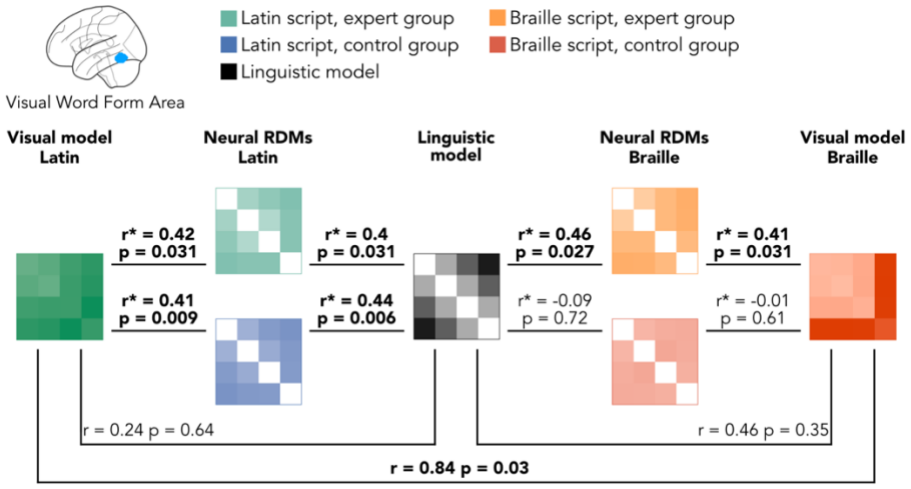

Figure 3-2 – Correlations between neural and model RDMs in the Visual Word Form Ara (VWFA). Each correlation between the neural matrices and a model was made as a partial correlation, controlling for the other model. As an example, in the correlations between the neural RDM of the Latin script in one group and the linguistic model, we partialed out the correlation between the neural RDM and visual model for the Latin script. In the case of neural RDM and visual model correlations, we partialed out the correlation between the neural RDM and the linguistic model. Correlations between models, not partial, show a limited correlation between linguistic and visual explanations of the neural organization. Partial correlations are represented by r*. Significant correlations are represented in bold.

Table 3-1 – Activation peaks for all the ROIs extracted from the functional localizer experiment

Region	Group	Hemisphe	Contrast	Subject	Z-score	Cluster size	MNI coordinates
VWFA	Experts	Left	FW > SFW	Exp-01	5.04	24	-47 -60 -13
				Exp-02	4.78	9	-52 -47 -18
				Exp-03	6.57	83	-52 -55 -18
				Exp-04	5.90	159	-44 -63 -13
				Exp-05	4.31	29	-44 -60 -15
				Exp-06	5.27	27	-44 -60 -13
			BW > SBW	Exp-01	5.09	113	-49 -63 -13
				Exp-02	5.23	92	-55 -47 -21
				Exp-03	NA	NA	NA
				Exp-04	3.50	7	-39 -60 -5
				Exp-05	4.23	37	-42 -57 -8
				Exp-06	5.39	79	-47 -63 -15

	Control	Left	FW > SFW	Ctr-01	Inf	536	-44 -66 -13
				Ctr-02	4.38	24	-49 -60 -10
				Ctr-03	6.06	203	-47 -55 -8
				Ctr-04	Inf	212	-47 -47 -20
				Ctr-05	6.80	413	-44 -57 -20
				Ctr-06	6.16	357	-47 -50 -23
				Ctr-07	3.71	4	-48 -65 -18
				Ctr-08	6.82	94	-47 -50 -23
				Ctr-09	3.51	24	-42 -52 -10
				Ctr-10	Inf	849	-52 -57 -15
				Ctr-11	Inf	510	-55 -57 -15
				Ctr-12	5.33	226	-55 -50 -13
	Control	Left	BW > SBW	Ctr-01	NA	NA	NA
				Ctr-02	NA	NA	NA
				Ctr-03	NA	NA	NA
				Ctr-04	NA	NA	NA
				Ctr-05	3.33	5	-47 -52 -21
				Ctr-06	NA	NA	NA
				Ctr-07	NA	NA	NA
				Ctr-08	NA	NA	NA
				Ctr-09	NA	NA	NA
				Ctr-10	NA	NA	NA
				Ctr-11	NA	NA	NA
				Ctr-12	NA	NA	NA
LOC	Experts	Left	LD > SLD	Exp-01	5.92	45	-39 -76 -15
				Exp-02	Inf	864	-44 -65 -13
				Exp-03	4.85	406	-42 -78 -2
				Exp-04	6.44	911	-49 -78 -3
				Exp-05	7.59	461	-49 -76 -10
				Exp-06	Inf	1355	-49 -78 -3
		Right	LD > SLD	Exp-01	5.68	100	39 -68 -10
				Exp-02	Inf	284	47 -78 0
				Exp-03	7.86	395	44 -78 -8
				Exp-04	6.26	335	37 -70 -13
				Exp-05	7.37	227	49 -73 -10
				Exp-06	Inf	669	49 -68 -13
	Control	Left	LD > SLD	Ctr-01	7.67	943	-49 -76 0
				Ctr-02	Inf	2176	-39 -73 -10
				Ctr-03	Inf	857	-47 -73 -5
				Ctr-04	Inf	717	-41 -78 -5
				Ctr-05	Inf	1147	-49 -78 -5
				Ctr-06	Inf	909	-47 -65 -13
				Ctr-07	Inf	446	-42 -75 -10
				Ctr-08	NA	NA	NA
				Ctr-09	NA	NA	NA
				Ctr-10	Inf	1043	-49 -73 -10

				Ctr-11	6.89	200	-52 -75 0
				Ctr-12	Inf	634	-47 -73 -13
		Right	LD > SLD	Ctr-01	6.83	613	49 -65 -5
				Ctr-02	Inf	2391	44 -65 -10
				Ctr-03	Inf	1097	44 -70 -10
				Ctr-04	6.55	81	44 -73 -7
				Ctr-05	Inf	327	44 -65 -13
				Ctr-06	6.73	539	49 -68 -7
				Ctr-07	4.90	36	49 -68 -10
				Ctr-08	5.20	71	39 -68 -5
				Ctr-09	NA	NA	NA
				Ctr-10	5.58	347	44 -68 -13
				Ctr-11	6.96	68	39 -76 -13
				Ctr-12	Inf	974	39 -73 -10

Figure 3-1 – Activation peaks for all the ROIs extracted from the functional localizer experiment. We localized VWFA individually in each expert, by contrasting the activation for intact Latin words over scrambled Latin words. Additionally, we localized activity for Braille by contrasting activation for intact Braille stimuli over scrambled ones. Bilateral Lateral Occipital (LO) was localized through the contrast of Line Drawing over scrambled Line Drawings. Details of the methods are available in the *Definition of the Regions of Interest (ROIs)* section

Table 3-2 – Post-hoc t-tests in Visual Word Form Area (VWFA)

Script	Effect	Comparison	Comparison contrasted	Mean Difference	SE	t	Cohen's d	p_{bonf}
Latin	Condition	RW-PW	RW-NW	-0.146	0.045	-3.266	-0.933	0.024
			RW-FS	-0.222	0.045	-4.976	-1.421	< .001
			PW-NW	0.024	0.045	0.544	0.155	1
			PW-FS	-0.181	0.045	-4.043	-1.155	0.002
			NW-FS	-0.174	0.045	-3.888	-1.11	0.003
		RW-NW	RW-FS	-0.076	0.045	-1.711	-0.488	1
			PW-NW	0.17	0.045	3.81	1.088	0.004
			PW-FS	-0.035	0.045	-0.778	-0.222	1
			NW-FS	-0.028	0.045	-0.622	-0.178	1
		RW-FS	PW-NW	0.247	0.045	5.52	1.576	< .001
			PW-FS	0.042	0.045	0.933	0.266	1
			NW-FS	0.049	0.045	1.089	0.311	1
		PW-NW	PW-FS	-0.205	0.045	-4.587	-1.31	< .001
			NW-FS	-0.198	0.045	-4.432	-1.266	< .001
		PW-FS	NW-FS	0.007	0.045	0.156	0.044	1
Braille	Grp*Cond.	experts, RW-PW	novices, RW-PW	0.049	0.066	0.738	0.369	1
			experts, RW-NW	-0.208	0.069	-3.026	-1.581	0.22
			novices, RW-NW	6×10^{-16}	0.066	10×10^{-15}	5×10^{-15}	1
			experts, RW-FS	-0.319	0.069	-4.639	-2.424	< .001
			novices, RW-FS	-0.014	0.066	-0.211	-0.105	1
			experts, PW-NW	-0.208	0.069	-3.026	-1.581	0.22

			novices, PW-NW	-0.042	0.066	-0.632	-0.316	1
			experts, PW-FS	-0.306	0.069	-4.438	-2.319	0.002
			novices, PW-FS	-0.007	0.066	-0.105	-0.053	1
			experts, NW-FS	-0.194	0.069	-2.824	-1.476	0.395
			novices, NW-FS	0.014	0.066	0.211	0.105	1
		novices, RW-PW	experts, RW-NW	-0.257	0.066	-3.9	-1.95	0.013
			novices, RW-NW	-0.049	0.049	-0.998	-0.369	1
			experts, RW-FS	-0.368	0.066	-5.587	-2.793	< .001
			novices, RW-FS	-0.062	0.049	-1.284	-0.474	1
			experts, PW-NW	-0.257	0.066	-3.9	-1.95	0.013
			novices, PW-NW	-0.09	0.049	-1.854	-0.685	1
			experts, PW-FS	-0.354	0.066	-5.376	-2.688	< .001
			novices, PW-FS	-0.056	0.049	-1.141	-0.422	1
			experts, NW-FS	-0.243	0.066	-3.689	-1.845	0.027
			novices, NW-FS	-0.035	0.049	-0.713	-0.264	1
		experts, RW-NW	novices, RW-NW	0.208	0.066	3.162	1.581	0.145
			experts, RW-FS	-0.111	0.069	-1.614	-0.843	1
			novices, RW-FS	0.194	0.066	2.951	1.476	0.272
			experts, PW-NW	3×10^{-17}	0.069	4×10^{-16}	0	1
			novices, PW-NW	0.167	0.066	2.53	1.265	0.879
			experts, PW-FS	-0.097	0.069	-1.412	-0.738	1
			novices, PW-FS	0.201	0.066	3.057	1.528	0.199
			experts, NW-FS	0.014	0.069	0.202	0.105	1
			novices, NW-FS	0.222	0.066	3.373	1.687	0.075
		novices, RW-NW	experts, RW-FS	-0.319	0.066	-4.849	-2.424	< .001
			novices, RW-FS	-0.014	0.049	-0.285	-0.105	1
			experts, PW-NW	-0.208	0.066	-3.162	-1.581	0.145
			novices, PW-NW	-0.042	0.049	-0.856	-0.316	1
			experts, PW-FS	-0.306	0.066	-4.638	-2.319	< .001
			novices, PW-FS	-0.007	0.049	-0.143	-0.053	1
			experts, NW-FS	-0.194	0.066	-2.951	-1.476	0.272
			novices, NW-FS	0.014	0.049	0.285	0.105	1
		experts, RW-FS	novices, RW-FS	0.306	0.066	4.638	2.319	< .001
			experts, PW-NW	0.111	0.069	1.614	0.843	1
			novices, PW-NW	0.278	0.066	4.216	2.108	0.004
			experts, PW-FS	0.014	0.069	0.202	0.105	1
			novices, PW-FS	0.313	0.066	4.743	2.372	< .001
			experts, NW-FS	0.125	0.069	1.815	0.949	1
			novices, NW-FS	0.333	0.066	5.06	2.53	< .001
		novices, RW-FS	experts, PW-NW	-0.194	0.066	-2.951	-1.476	0.272
			novices, PW-NW	-0.028	0.049	-0.571	-0.211	1
			experts, PW-FS	-0.292	0.066	-4.427	-2.214	0.002
			novices, PW-FS	0.007	0.049	0.143	0.053	1
			experts, NW-FS	-0.181	0.066	-2.741	-1.37	0.496
			novices, NW-FS	0.028	0.049	0.571	0.211	1
		experts, PW-NW	novices, PW-NW	0.167	0.066	2.53	1.265	0.879
			experts, PW-FS	-0.097	0.069	-1.412	-0.738	1
			novices, PW-FS	0.201	0.066	3.057	1.528	0.199
			experts, NW-FS	0.014	0.069	0.202	0.105	1
			novices, NW-FS	0.222	0.066	3.373	1.687	0.075
		novices, PW-NW	experts, PW-FS	-0.264	0.066	-4.006	-2.003	0.009
			novices, PW-FS	0.035	0.049	0.713	0.264	1
			experts, NW-FS	-0.153	0.066	-2.319	-1.16	1
			novices, NW-FS	0.056	0.049	1.141	0.422	1
		experts, PW-FS	novices, PW-FS	0.299	0.066	4.533	2.266	0.001
			experts, NW-FS	0.111	0.069	1.614	0.843	1
			novices, NW-FS	0.319	0.066	4.849	2.424	< .001
		novices, PW-FS	experts, NW-FS	-0.187	0.066	-2.846	-1.423	0.368
			novices, NW-FS	0.021	0.049	0.428	0.158	1
		experts, NW-FS	novices, NW-FS	0.208	0.066	3.162	1.581	0.145

Figure 3-2 – Post-hoc t-tests in Visual Word Form Area (VWFA). For each effect in the repeated measures ANOVA (rmANOVA), we contrasted the different comparisons to identify the causes of the interaction effects. In Latin script, results are averaged over the levels of 'group'. P-values are adjusted for comparing a family of 15. In Braille script, p-values are adjusted for comparing a family of 66.

Table 4-1 – Statistical tests in left-LO

Analysis	Conditions	Statistical test	Result	DF	p-value	Effect size
Pairwise decoding (average)	Latin, experts	t-test	8.58	5	< 0.001	D = 3.5
	Latin, controls	t-test	8.57	9	< 0.001	D = 2.7
	Braille, experts	t-test	4.48	5	0.006	D = 1.8
	Braille, controls	t-test	1.45	9	0.1	D = 0.5
	Latin, experts – Latin, controls	t-test	1.05	11	0.31	D = 0.5
	Latin, experts – Braille, experts	Paired t-test	3.26	5	0.03	D = 1.9
	Latin, experts – Braille, controls	t-test	7.39	11	< 0.001	D = 4.4
	Latin, controls – Braille, experts	t-test	-2.51	11	0.034	D = 1.2
	Latin, controls – Braille, controls	Paired t-test	8.83	10	< 0.001	D = 2.9
	Braille, experts – Braille, controls	t-test	3.50	7	0.017	D = 2
Within Braille ANOVA	Effect of group	rmANOVA	16.01	1	0.001	$\eta^2_p = 0.53$
	Effect of conditions	rmANOVA	1.72	5	0.14	$\eta^2_p = 0.11$
	Interaction group and condition	rmANOVA	0.72	5	0.61	$\eta^2_p = 0.05$
Within Latin ANOVA	Effect of group	rmANOVA	1.06	1	0.32	$\eta^2_p = 0.07$
	Effect of conditions	rmANOVA	35.61	5	< 0.001	$\eta^2_p = 0.72$
	Interaction group and condition	rmANOVA	2.94	5	0.02	$\eta^2_p = 0.17$
RSA	Latin, experts – Latin, controls	Pearson's R	0.75	NA	< 0.001	NA
	Latin, experts – Braille, experts	Pearson's R	0.56	NA	< 0.001	NA
	Latin, experts – Braille, controls	Pearson's R	0.19	NA	0.07	NA
	Latin, controls – Braille, experts	Pearson's R	0.55	NA	< 0.001	NA
	Latin, controls – Braille, controls	Pearson's R	0.14	NA	0.13	NA
	Braille, experts – Braille, controls	Pearson's R	0.14	NA	0.15	NA
	Latin, experts – Model	Pearson's R	0.51	NA	0.002	NA
	Braille, experts – Model	Pearson's R	0.34	NA	0.052	NA
	Latin, controls – Model	Pearson's R	0.61	NA	< 0.001	NA
	Braille, controls – Model	Pearson's R	-0.06	NA	0.68	NA
Cross-script decoding	Average	t-test	1.56	5	0.16	D = 0.6
	Real words – Pseudo-words	t-test	4.39	5	0.025	D = 1.8
	Real words – non-words	t-test	0.37	5	0.42	D = 0.2
	Real words – fake script	t-test	0.68	5	0.16	D = 0.7
	Pseudo-words – non-words	t-test	-0.75	5	0.76	D = 0.3
	Pseudo-words – fake script	t-test	1.38	5	0.16	D = 0.6
	Non-words – fake script	t-test	1.57	5	0.16	D = 0.6

Figure 4-1 – Statistical tests in left-LO. Full list of statistical tests performed over the multivariate results of l-PosTemp. Each analysis is detailed with the conditions tested and the statistical test. Statistical tests detail the result, the degrees of freedom (DF), the p-value, and a measure of the effect size. For t-test, effect size was measured through Cohen's D. For repeated measures ANOVA (rmANOVA), effect size was measured through partial eta-squared. In the case of Pearson's correlation (R), the value of the correlation is to be taken as the measure of the effect size

Table 4-2 – Post-hoc t-tests in left-LO

Script	Effect	Comparison	Comparison contrasted	Mean Difference	SE	t	Cohen's d	p_{bonf}
Latin	Condition	RW-PW	RW-NW	-0.253	0.04	-6.349	-1.877	< .001
			RW-FS	-0.394	0.04	-9.907	-2.929	< .001
			PW-NW	-0.078	0.04	-1.953	-0.578	0.821
			PW-FS	-0.383	0.04	-9.628	-2.847	< .001
			NW-FS	-0.353	0.04	-8.86	-2.62	< .001
		RW-NW	RW-FS	-0.142	0.04	-3.558	-1.052	0.01
			PW-NW	0.175	0.04	4.395	1.3	< .001
			PW-FS	-0.131	0.04	-3.279	-0.97	0.024
			NW-FS	-0.1	0.04	-2.512	-0.743	0.215
		RW-FS	PW-NW	0.317	0.04	7.953	2.352	< .001
			PW-FS	0.011	0.04	0.279	0.083	1
			NW-FS	0.042	0.04	1.047	0.309	1
		PW-NW	PW-FS	-0.306	0.04	-7.674	-2.269	< .001
			NW-FS	-0.275	0.04	-6.907	-2.042	< .001
		PW-FS	NW-FS	0.031	0.04	0.767	0.227	1
Braille	Grp*Cond.	experts, RW-PW	novices, RW-PW	0.042	0.073	0.57	0.295	1
			experts, RW-NW	-0.125	0.081	-1.54	-0.884	1
			novices, RW-NW	0.058	0.073	0.799	0.412	1
			experts, RW-FS	-0.139	0.081	-1.711	-0.982	1
			novices, RW-FS	0.025	0.073	0.342	0.177	1
			experts, PW-NW	-0.042	0.081	-0.513	-0.295	1
			novices, PW-NW	0.017	0.073	0.228	0.118	1
			experts, PW-FS	-0.208	0.081	-2.566	-1.473	0.82
			novices, PW-FS	-0.033	0.073	-0.456	-0.236	1
			experts, NW-FS	-0.097	0.081	-1.197	-0.687	1
			novices, NW-FS	0.017	0.073	0.228	0.118	1
		novices, RW-PW	experts, RW-NW	-0.167	0.073	-2.282	-1.178	1
			novices, RW-NW	0.017	0.063	0.265	0.118	1
			experts, RW-FS	-0.181	0.073	-2.472	-1.277	1
			novices, RW-FS	-0.017	0.063	-0.265	-0.118	1
			experts, PW-NW	-0.083	0.073	-1.141	-0.589	1
			novices, PW-NW	-0.025	0.063	-0.398	-0.177	1
			experts, PW-FS	-0.25	0.073	-3.423	-1.768	0.063
			novices, PW-FS	-0.075	0.063	-1.193	-0.53	1
			experts, NW-FS	-0.139	0.073	-1.902	-0.982	1
			novices, NW-FS	-0.025	0.063	-0.398	-0.177	1
		experts, RW-NW	novices, RW-NW	0.183	0.073	2.51	1.296	0.923
			experts, RW-FS	-0.014	0.081	-0.171	-0.098	1
			novices, RW-FS	0.15	0.073	2.054	1.061	1
			experts, PW-NW	0.083	0.081	1.026	0.589	1
			novices, PW-NW	0.142	0.073	1.94	1.002	1
			experts, PW-FS	-0.083	0.081	-1.026	-0.589	1
			novices, PW-FS	0.092	0.073	1.255	0.648	1
			experts, NW-FS	0.028	0.081	0.342	0.196	1
			novices, NW-FS	0.142	0.073	1.94	1.002	1
		novices, RW-NW	experts, RW-FS	-0.197	0.073	-2.7	-1.394	0.553

			novices, RW-FS	-0.033	0.063	-0.53	-0.236	1
			experts, PW-NW	-0.1	0.073	-1.369	-0.707	1
			novices, PW-NW	-0.042	0.063	-0.663	-0.295	1
			experts, PW-FS	-0.267	0.073	-3.651	-1.885	0.03
			novices, PW-FS	-0.092	0.063	-1.458	-0.648	1
			experts, NW-FS	-0.156	0.073	-2.13	-1.1	1
			novices, NW-FS	-0.042	0.063	-0.663	-0.295	1
		experts, RW-FS	novices, RW-FS	0.164	0.073	2.244	1.159	1
			experts, PW-NW	0.097	0.081	1.197	0.687	1
			novices, PW-NW	0.156	0.073	2.13	1.1	1
			experts, PW-FS	-0.069	0.081	-0.855	-0.491	1
			novices, PW-FS	0.106	0.073	1.445	0.746	1
			experts, NW-FS	0.042	0.081	0.513	0.295	1
			novices, NW-FS	0.156	0.073	2.13	1.1	1
		novices, RW-FS	experts, PW-NW	-0.067	0.073	-0.913	-0.471	1
			novices, PW-NW	-0.008	0.063	-0.133	-0.059	1
			experts, PW-FS	-0.233	0.073	-3.195	-1.65	0.13
			novices, PW-FS	-0.058	0.063	-0.928	-0.412	1
			experts, NW-FS	-0.122	0.073	-1.673	-0.864	1
			novices, NW-FS	-0.008	0.063	-0.133	-0.059	1
		experts, PW-NW	novices, PW-NW	0.058	0.073	0.799	0.412	1
			experts, PW-FS	-0.167	0.081	-2.053	-1.178	1
			novices, PW-FS	0.008	0.073	0.114	0.059	1
			experts, NW-FS	-0.056	0.081	-0.684	-0.393	1
			novices, NW-FS	0.058	0.073	0.799	0.412	1
		novices, PW-NW	experts, PW-FS	-0.225	0.073	-3.081	-1.591	0.184
			novices, PW-FS	-0.05	0.063	-0.795	-0.354	1
			experts, NW-FS	-0.114	0.073	-1.559	-0.805	1
			novices, NW-FS	1.1×10^{-16}	0.063	2×10^{-15}	2×10^{-15}	1
		experts, PW-FS	novices, PW-FS	0.175	0.073	2.396	1.237	1
			experts, NW-FS	0.111	0.081	1.369	0.786	1
			novices, NW-FS	0.225	0.073	3.081	1.591	0.184
		novices, PW-FS	experts, NW-FS	-0.064	0.073	-0.875	-0.452	1
			novices, NW-FS	0.05	0.063	0.795	0.354	1
		experts, NW-FS	novices, NW-FS	0.114	0.073	1.559	0.805	1

Figure 4-2 – Post-hoc t-tests in left-LO. For each effect in the repeated measures ANOVA (rmANOVA), we contrasted the different comparisons to identify the causes of the interaction effects. In Latin script, results are averaged over the levels of 'group'. P-values are adjusted for comparing a family of 15. In Braille script, p-values are adjusted for comparing a family of 66.

Table 4-3 – Statistical tests in right-LO

Analysis	Conditions	Statistical test	Result	DF	p-value	Effect size
Pairwise decoding (average)	Latin, experts	t-test	9.63	5	< 0.001	D = 3.9
	Latin, controls	t-test	5.87	10	< 0.001	D = 1.7
	Braille, experts	t-test	3.88	5	0.012	D = 1.5
	Braille, controls	t-test	1.01	10	0.21	D = 0.3
	Latin, experts – Latin, controls	t-test	0.45	15	0.66	D = 0.2
	Latin, experts – Braille, experts	Paired t-test	1.85	5	0.17	D = 1.1
	Latin, experts – Braille, controls	t-test	6.64	12	< 0.001	D = 3.2
	Latin, controls – Braille, experts	t-test	-1.25	13	0.25	D = 0.6

	Latin, controls – Braille, controls	Paired t-test	6.57	10	< 0.001	D = 1.8
	Braille, experts – Braille, controls	t-test	2.92	8	0.032	D = 1.6
Within Braille ANOVA	Effect of group	rmANOVA	10.23	1	0.006	η^2_p = 0.41
	Effect of conditions	rmANOVA	1.89	5	0.11	η^2_p = 0.11
	Interaction group and condition	rmANOVA	1.41	5	0.23	η^2_p = 0.09
Within Latin ANOVA	Effect of group	rmANOVA	0.14	1	0.72	η^2_p = 0.01
	Effect of conditions	rmANOVA	35.70	5	< 0.001	η^2_p = 0.70
	Interaction group and condition	rmANOVA	0.90	5	0.49	η^2_p = 0.06
RSA	Latin, experts – Latin, controls	Pearson's R	0.83	NA	< 0.001	NA
	Latin, experts – Braille, experts	Pearson's R	0.65	NA	< 0.001	NA
	Latin, experts – Braille, controls	Pearson's R	0.06	NA	0.36	NA
	Latin, controls – Braille, experts	Pearson's R	0.64	NA	< 0.001	NA
	Latin, controls – Braille, controls	Pearson's R	0.01	NA	0.52	NA
	Braille, experts – Braille, controls	Pearson's R	0.04	NA	0.44	NA
	Latin, experts – Model	Pearson's R	0.64	NA	< 0.001	NA
	Braille, experts – Model	Pearson's R	0.53	NA	0.002	NA
	Latin, controls – Model	Pearson's R	0.44	NA	< 0.001	NA
	Braille, controls – Model	Pearson's R	0.04	NA	0.46	NA
Cross-script decoding	Average	t-test	2.35	5	0.1	D = 1
	Real words – Pseudo-words	t-test	-0.36	5	0.63	D = 0.1
	Real words – non-words	t-test	-0.92	5	0.1	D = 0.4
	Real words – fake script	t-test	1.83	5	0.1	D = 0.7
	Pseudo-words – non-words	t-test	1.75	5	0.1	D = 0.7
	Pseudo-words – fake script	t-test	1.94	5	0.09	D = 0.8
	Non-words – fake script	t-test	2.28	5	0.1	D = 0.9

Figure 4-3 – Statistical tests in right-LO. Full list of statistical tests performed over the multivariate results of l-PosTemp. Each analysis is detailed with the conditions tested and the statistical test. Statistical tests detail the result, the degrees of freedom (DF), the p-value, and a measure of the effect size. For t-test, effect size was measured through Cohen's D. For repeated measures ANOVA (rmANOVA), effect size was measured through partial eta-squared. In the case of Pearson's correlation (R), the value of the correlation is to be taken as the measure of the effect size.

Table 4-4 – Post-hoc tests in right-LO

Script	Effect	Comparison	Comparison contrasted	Mean Difference	SE	t	Cohen's d	p_{bonf}
Latin	Condition	RW-PW	RW-NW	-0.244	0.042	-5.865	-1.578	< .001
			RW-FS	-0.436	0.042	-10.472	-2.817	< .001
			PW-NW	-0.16	0.042	-3.849	-1.036	0.004
			PW-FS	-0.438	0.042	-10.518	-2.829	< .001
			NW-FS	-0.387	0.042	-9.29	-2.499	< .001
		RW-NW	RW-FS	-0.192	0.042	-4.607	-1.239	< .001
			PW-NW	0.084	0.042	2.016	0.542	0.711
			PW-FS	-0.194	0.042	-4.653	-1.252	< .001
			NW-FS	-0.143	0.042	-3.425	-0.921	0.015
		RW-FS	PW-NW	0.276	0.042	6.623	1.782	< .001
			PW-FS	-0.002	0.042	-0.045	-0.012	1

			NW-FS	0.049	0.042	1.182	0.318	1
		PW-NW	PW-FS	-0.278	0.042	-6.668	-1.794	< .001
			NW-FS	-0.227	0.042	-5.441	-1.464	< .001
		PW-FS	NW-FS	0.051	0.042	1.228	0.33	1
Braille	Grp*Cond.	experts, RW-PW	novices, RW-PW	0.016	0.076	0.217	0.11	1
			experts, RW-NW	-0.056	0.078	-0.716	-0.373	1
			novices, RW-NW	0.077	0.076	1.019	0.517	1
			experts, RW-FS	-0.194	0.078	-2.506	-1.306	0.95
			novices, RW-FS	0.009	0.076	0.117	0.059	1
			experts, PW-NW	-0.042	0.078	-0.537	-0.28	1
			novices, PW-NW	0.1	0.076	1.32	0.67	1
			experts, PW-FS	-0.181	0.078	-2.327	-1.212	1
			novices, PW-FS	0.062	0.076	0.819	0.415	1
			experts, NW-FS	-0.069	0.078	-0.895	-0.466	1
			novices, NW-FS	0.016	0.076	0.217	0.11	1
		novices, RW-PW	experts, RW-NW	-0.072	0.076	-0.952	-0.483	1
			novices, RW-NW	0.061	0.057	1.057	0.407	1
			experts, RW-FS	-0.211	0.076	-2.79	-1.416	0.439
			novices, RW-FS	-0.008	0.057	-0.132	-0.051	1
			experts, PW-NW	-0.058	0.076	-0.768	-0.39	1
			novices, PW-NW	0.083	0.057	1.454	0.56	1
			experts, PW-FS	-0.197	0.076	-2.606	-1.323	0.726
			novices, PW-FS	0.045	0.057	0.793	0.305	1
			experts, NW-FS	-0.086	0.076	-1.136	-0.576	1
			novices, NW-FS	3.5×10^{-17}	0.057	6×10^{-16}	1×10^{-15}	1
		experts, RW-NW	novices, RW-NW	0.133	0.076	1.754	0.89	1
			experts, RW-FS	-0.139	0.078	-1.79	-0.933	1
			novices, RW-FS	0.064	0.076	0.852	0.432	1
			experts, PW-NW	0.014	0.078	0.179	0.093	1
			novices, PW-NW	0.155	0.076	2.055	1.043	1
			experts, PW-FS	-0.125	0.078	-1.611	-0.839	1
			novices, PW-FS	0.117	0.076	1.554	0.788	1
			experts, NW-FS	-0.014	0.078	-0.179	-0.093	1
			novices, NW-FS	0.072	0.076	0.952	0.483	1
		novices, PW-NW	experts, RW-FS	-0.271	0.076	-3.591	-1.823	0.038
			novices, RW-FS	-0.068	0.057	-1.19	-0.458	1
			experts, PW-NW	-0.119	0.076	-1.57	-0.797	1
			novices, PW-NW	0.023	0.057	0.397	0.153	1
			experts, PW-FS	-0.258	0.076	-3.408	-1.729	0.069
			novices, PW-FS	-0.015	0.057	-0.264	-0.102	1
			experts, NW-FS	-0.146	0.076	-1.938	-0.983	1
			novices, NW-FS	-0.061	0.057	-1.057	-0.407	1
		experts, RW-FS	novices, RW-FS	0.203	0.076	2.689	1.365	0.579
			experts, PW-NW	0.153	0.078	1.969	1.026	1
			novices, PW-NW	0.294	0.076	3.892	1.975	0.014
			experts, PW-FS	0.014	0.078	0.179	0.093	1
			novices, PW-FS	0.256	0.076	3.391	1.721	0.073
			experts, NW-FS	0.125	0.078	1.611	0.839	1
			novices, NW-FS	0.211	0.076	2.79	1.416	0.439
		novices, RW-FS	experts, PW-NW	-0.051	0.076	-0.668	-0.339	1
			novices, PW-NW	0.091	0.057	1.586	0.61	1
			experts, PW-FS	-0.189	0.076	-2.506	-1.272	0.946
			novices, PW-FS	0.053	0.057	0.925	0.356	1
			experts, NW-FS	-0.078	0.076	-1.036	-0.526	1
			novices, NW-FS	0.008	0.057	0.132	0.051	1
		experts, PW-NW	novices, PW-NW	0.141	0.076	1.871	0.95	1

			experts, PW-FS	-0.139	0.078	-1.79	-0.933	1
			novices, PW-FS	0.104	0.076	1.37	0.695	1
			experts, NW-FS	-0.028	0.078	-0.358	-0.187	1
			novices, NW-FS	0.058	0.076	0.768	0.39	1
		novices, PW-NW	experts, PW-FS	-0.28	0.076	-3.708	-1.882	0.026
			novices, PW-FS	-0.038	0.057	-0.661	-0.254	1
			experts, NW-FS	-0.169	0.076	-2.238	-1.136	1
			novices, NW-FS	-0.083	0.057	-1.454	-0.56	1
		experts, PW-FS	novices, PW-FS	0.242	0.076	3.207	1.628	0.129
			experts, NW-FS	0.111	0.078	1.432	0.746	1
			novices, NW-FS	0.197	0.076	2.606	1.323	0.726
		novices, PW-FS	experts, NW-FS	-0.131	0.076	-1.737	-0.882	1
			novices, NW-FS	-0.045	0.057	-0.793	-0.305	1
		experts, NW-FS	novices, NW-FS	0.086	0.076	1.136	0.576	1

Figure 4-4 – Post-hoc t-tests in right-LO. For each effect in the repeated measures ANOVA (rmANOVA), we contrasted the different comparisons to identify the causes of the interaction effects. In Latin script, results are averaged over the levels of 'group'. P-values are adjusted for comparing a family of 15. In Braille script, p-values are adjusted for comparing a family of 66.

Table 5-1 – Statistical tests in V1

Analysis	Conditions	Statistical test	Result	DF	p-value	Effect size
Pairwise decoding (average)	Latin, experts	t-test	4.35	5	0.018	D = 1.8
	Latin, controls	t-test	3.15	10	0.018	D = 0.9
	Braille, experts	t-test	3.93	5	0.018	D = 1.6
	Braille, controls	t-test	2.45	10	0.28	D = 0.7
	Latin, experts – Latin, controls	t-test	0.71	13	0.53	D = 0.3
	Latin, experts – Braille, experts	Paired t-test	1.60	5	0.21	D = 0.6
	Latin, experts – Braille, controls	t-test	3.47	6	0.027	D = 2.2
	Latin, controls – Braille, experts	t-test	0.23	15	0.82	D = 0.1
	Latin, controls – Braille, controls	Paired t-test	3.01	10	0.027	D = 0.7
	Braille, experts – Braille, controls	t-test	2.78	6	0.04	D = 1.7
Within Braille ANOVA	Effect of group	rmANOVA	11.12	1	0.005	η^2_p = 0.43
	Effect of conditions	rmANOVA	4.05	5	0.003	η^2_p = 0.21
	Interaction group and condition	rmANOVA	2.33	5	0.051	η^2_p = 0.13
Within Latin ANOVA	Effect of group	rmANOVA	0.41	1	0.53	η^2_p = 0.03
	Effect of conditions	rmANOVA	11.09	5	< 0.001	η^2_p = 0.43
	Interaction group and condition	rmANOVA	1.32	5	0.26	η^2_p = 0.08
RSA	Latin, experts – Latin, controls	Pearson's R	0.56	NA	< 0.001	NA

	Latin, experts – Braille, experts	Pearson's R	0.50	NA	< 0.001	NA
	Latin, experts – Braille, controls	Pearson's R	0.20	NA	0.06	NA
	Latin, controls – Braille, experts	Pearson's R	0.49	NA	< 0.001	NA
	Latin, controls – Braille, controls	Pearson's R	0.22	NA	0.02	NA
	Braille, experts – Braille, controls	Pearson's R	0.18	NA	0.07	NA
	Latin, experts – Model	Pearson's R	0.68	NA	< 0.001	NA
	Braille, experts – Model	Pearson's R	0.25	NA	0.09	NA
	Latin, controls – Model	Pearson's R	0.28	NA	0.025	NA
	Braille, controls – Model	Pearson's R	0.14	NA	0.15	NA
Cross-script decoding	Average	t-test	-0.36	5	0.86	D = 0.1
	Real words – Pseudo-words	t-test	-0.79	5	0.86	D = 0.3
	Real words – non-words	t-test	-1.21	5	0.86	D = 0.5
	Real words – fake script	t-test	0.73	5	0.71	D = 0.3
	Pseudo-words – non-words	t-test	-1.03	5	0.86	D = 0.4
	Pseudo-words – fake script	t-test	0.54	5	0.71	D = 0.2
	Non-words – fake script	t-test	0.88	5	0.71	D = 0.3

Figure 5-1 – Statistical tests in V1. Full list of statistical tests performed over the multivariate results of l-PosTemp. Each analysis is detailed with the conditions tested and the statistical test. Statistical tests detail the result, the degrees of freedom (DF), the p-value, and a measure of the effect size. For t-test, effect size was measured through Cohen's D. For repeated measures ANOVA (rmANOVA), effect size was measured through partial eta-squared. In the case of Pearson's correlation (R), the value of the correlation is to be taken as the measure of the effect size

Table 5-2 – Post-hoc t-tests in V1

Script	Effect	Comparison	Comparison contrasted	Mean Difference	SE	t	Cohen's d	p_{bonf}
Latin	Condition	RW-PW	RW-NW	-0.197	0.043	-4.592	-1.066	< .001
			RW-FS	-0.292	0.043	-6.8	-1.579	< .001
			PW-NW	-0.165	0.043	-3.842	-0.892	0.004
			PW-FS	-0.255	0.043	-5.947	-1.38	< .001
			NW-FS	-0.18	0.043	-4.195	-0.974	0.001
		RW-NW	RW-FS	-0.095	0.043	-2.208	-0.513	0.455
			PW-NW	0.032	0.043	0.751	0.174	1
			PW-FS	-0.058	0.043	-1.354	-0.314	1
			NW-FS	0.017	0.043	0.397	0.092	1
		RW-FS	PW-NW	0.127	0.043	2.959	0.687	0.062
			PW-FS	0.037	0.043	0.854	0.198	1
			NW-FS	0.112	0.043	2.605	0.605	0.166
		PW-NW	PW-FS	-0.09	0.043	-2.105	-0.489	0.58
			NW-FS	-0.015	0.043	-0.353	-0.082	1
		PW-FS	NW-FS	0.075	0.043	1.752	0.407	1
Braille	Grp*Cond.	experts, RW-PW	novices, RW-PW	-0.049	0.075	-0.658	-0.334	1
			experts, RW-NW	-0.111	0.084	-1.32	-0.754	1
			novices, RW-NW	-0.11	0.075	-1.469	-0.745	1

			experts, RW-FS	-0.292	0.084	-3.465	-1.979	0.058
			novices, RW-FS	-0.072	0.075	-0.962	-0.488	1
			experts, PW-NW	-0.111	0.084	-1.32	-0.754	1
			novices, PW-NW	-0.019	0.075	-0.253	-0.129	1
			experts, PW-FS	-0.264	0.084	-3.135	-1.791	0.162
			novices, PW-FS	-0.102	0.075	-1.367	-0.694	1
			experts, NW-FS	-0.319	0.084	-3.795	-2.168	0.02
			novices, NW-FS	-0.102	0.075	-1.367	-0.694	1
		novices, RW-PW	experts, RW-NW	-0.062	0.075	-0.827	-0.42	1
			novices, RW-NW	-0.061	0.062	-0.975	-0.411	1
			experts, RW-FS	-0.242	0.075	-3.241	-1.645	0.11
			novices, RW-FS	-0.023	0.062	-0.366	-0.154	1
			experts, PW-NW	-0.062	0.075	-0.827	-0.42	1
			novices, PW-NW	0.03	0.062	0.487	0.206	1
			experts, PW-FS	-0.215	0.075	-2.87	-1.456	0.338
			novices, PW-FS	-0.053	0.062	-0.853	-0.36	1
			experts, NW-FS	-0.27	0.075	-3.613	-1.833	0.033
			novices, NW-FS	-0.053	0.062	-0.853	-0.36	1
		experts, RW-NW	novices, RW-NW	0.001	0.075	0.017	0.009	1
			experts, RW-FS	-0.181	0.084	-2.145	-1.225	1
			novices, RW-FS	0.039	0.075	0.523	0.266	1
			experts, PW-NW	$3.6x10^{-17}$	0.084	$4.3×10^{-16}$	0	1
			novices, PW-NW	0.092	0.075	1.232	0.625	1
			experts, PW-FS	-0.153	0.084	-1.815	-1.037	1
			novices, PW-FS	0.009	0.075	0.118	0.06	1
			experts, NW-FS	-0.208	0.084	-2.475	-1.414	1
			novices, NW-FS	0.009	0.075	0.118	0.06	1
		novices, RW-NW	experts, RW-FS	-0.182	0.075	-2.431	-1.234	1
			novices, RW-FS	0.038	0.062	0.609	0.257	1
			experts, PW-NW	-0.001	0.075	-0.017	-0.009	1
			novices, PW-NW	0.091	0.062	1.462	0.617	1
			experts, PW-FS	-0.154	0.075	-2.06	-1.045	1
			novices, PW-FS	0.008	0.062	0.122	0.051	1
			experts, NW-FS	-0.21	0.075	-2.802	-1.422	0.41
			novices, NW-FS	0.008	0.062	0.122	0.051	1
		experts, RW-FS	novices, RW-FS	0.22	0.075	2.937	1.491	0.278
			experts, PW-NW	0.181	0.084	2.145	1.225	1
			novices, PW-NW	0.273	0.075	3.646	1.851	0.029
			experts, PW-FS	0.028	0.084	0.33	0.188	1
			novices, PW-FS	0.189	0.075	2.532	1.285	0.863
			experts, NW-FS	-0.028	0.084	-0.33	-0.188	1
			novices, NW-FS	0.189	0.075	2.532	1.285	0.863
		novices, RW-FS	experts, PW-NW	-0.039	0.075	-0.523	-0.266	1
			novices, PW-NW	0.053	0.062	0.853	0.36	1
			experts, PW-FS	-0.192	0.075	-2.566	-1.302	0.788
			novices, PW-FS	-0.03	0.062	-0.487	-0.206	1
			experts, NW-FS	-0.247	0.075	-3.309	-1.679	0.089
			novices, NW-FS	-0.03	0.062	-0.487	-0.206	1
		experts, PW-NW	novices, PW-NW	0.092	0.075	1.232	0.625	1
			experts, PW-FS	-0.153	0.084	-1.815	-1.037	1
			novices, PW-FS	0.009	0.075	0.118	0.06	1
			experts, NW-FS	-0.208	0.084	-2.475	-1.414	1
			novices, NW-FS	0.009	0.075	0.118	0.06	1
		novices, PW-NW	experts, PW-FS	-0.245	0.075	-3.275	-1.662	0.099
			novices, PW-FS	-0.083	0.062	-1.341	-0.565	1
			experts, NW-FS	-0.301	0.075	-4.018	-2.039	0.008
			novices, NW-FS	-0.083	0.062	-1.341	-0.565	1
		experts, PW-FS	novices, PW-FS	0.162	0.075	2.161	1.097	1
			experts, NW-FS	-0.056	0.084	-0.66	-0.377	1
			novices, NW-FS	0.162	0.075	2.161	1.097	1
		novices, PW-FS	experts, NW-FS	-0.217	0.075	-2.904	-1.474	0.306

			novices, NW-FS	6.9×10^{-18}	0.062	1×10^{-16}	0	1
		experts, NW-FS	novices, NW-FS	0.217	0.075	2.904	1.474	0.306

Figure 5-2 – Post-hoc t-tests in V1. For each effect in the repeated measures ANOVA (rmANOVA), we contrasted the different comparisons to identify the causes of the interaction effects. In Latin script, results are averaged over the levels of 'group'. P-values are adjusted for comparing a family of 15. In Braille script, p-values are adjusted for comparing a family of 66.

Table 6-1 – Statistical tests in left PosTemp

Analysis	Conditions	Statistical test	Result	DF	p-value	Effect size
Pairwise decoding (average)	Latin, experts	t-test	10.12	4	< 0.001	D = 4.5
	Latin, controls	t-test	5.35	9	< 0.001	D = 1.7
	Braille, experts	t-test	16.66	4	< 0.001	D = 7.4
	Braille, controls	t-test	0.92	9	0.24	D = 0.3
	Latin, experts – Latin, controls	t-test	1.14	12	0.30	D = 0.5
	Latin, experts – Braille, experts	Paired t-test	-1.08	4	0.34	D = 0.4
	Latin, experts – Braille, controls	t-test	6.96	10	< 0.001	D = 3.5
	Latin, controls – Braille, experts	t-test	1.90	12	0.11	D = 0.7
	Latin, controls – Braille, controls	Paired t-test	7.88	9	< 0.001	D = 1.5
	Braille, experts – Braille, controls	t-test	9.32	12	< 0.001	D = 4.2
Within Braille ANOVA	Effect of group	rmANOVA	58.42	1	< 0.001	η^2_p = 0.82
	Effect of conditions	rmANOVA	4.30	5	0.002	η^2_p = 0.25
	Interaction group and condition	rmANOVA	2.27	5	0.06	η^2_p = 0.15
Within Latin ANOVA	Effect of group	rmANOVA	0.83	1	0.38	η^2_p = 0.06
	Effect of conditions	rmANOVA	4.99	5	< 0.001	η^2_p = 0.28
	Interaction group and condition	rmANOVA	1.27	5	0.31	η^2_p = 0.09
RSA	Latin, experts – Latin, controls	Pearson's R	0.43	NA	< 0.001	NA
	Latin, experts – Braille, experts	Pearson's R	0.52	NA	< 0.001	NA
	Latin, experts – Braille, controls	Pearson's R	0.15	NA	0.13	NA
	Latin, controls – Braille, experts	Pearson's R	0.44	NA	< 0.001	NA
	Latin, controls – Braille, controls	Pearson's R	0.18	NA	0.06	NA
	Braille, experts – Braille, controls	Pearson's R	0.15	NA	0.13	NA
	Latin, experts – Model	Pearson's R	0.18	NA	0.21	NA
	Braille, experts – Model	Pearson's R	0.27	NA	0.13	NA
	Latin, controls – Model	Pearson's R	0.26	NA	0.06	NA
	Braille, controls – Model	Pearson's R	-0.06	NA	0.65	NA
Cross-script decoding	Average	t-test	4.08	4	0.025	D = 1.8
	Real words – Pseudo-words	t-test	2.24	4	0.052	D = 1
	Real words – non-words	t-test	4.49	4	0.025	D = 2
	Real words – fake script	t-test	0.29	4	0.39	D = 0.1

	Pseudo-words – non-words	t-test	3.64	4	0.025	D = 1.6
	Pseudo-words – fake script	t-test	2.69	4	0.04	D = 1.2
	Non-words – fake script	t-test	3.35	4	0.025	D = 1.5

Figure 6-1 – Statistical tests in l-PosTemp. Full list of statistical tests performed over the multivariate results of l-PosTemp. Each analysis is detailed with the conditions tested and the statistical test. Statistical tests detail the result, the degrees of freedom (DF), the p-value, and a measure of the effect size. For t-test, effect size was measured through Cohen's D. For repeated measures ANOVA (rmANOVA), effect size was measured through partial eta-squared. In the case of Pearson's correlation (R), the value of the correlation is to be taken as the measure of the effect size.

Table 6-2 – Post-hoc tests in left PosTemp

Script	Effect	Comparison	Comparison contrasted	Mean Difference	SE	t	Cohen's d	p_{bonf}
Latin	Condition	RW-PW	RW-NW	-0.221	0.053	-4.187	-1.281	0.001
			RW-FS	-0.183	0.053	-3.476	-1.063	0.014
			PW-NW	-0.121	0.053	-2.291	-0.701	0.378
			PW-FS	-0.179	0.053	-3.397	-1.039	0.017
			NW-FS	-0.217	0.053	-4.108	-1.257	0.002
		RW-NW	RW-FS	0.037	0.053	0.711	0.217	1
			PW-NW	0.1	0.053	1.896	0.58	0.936
			PW-FS	0.042	0.053	0.79	0.242	1
			NW-FS	0.004	0.053	0.079	0.024	1
		RW-FS	PW-NW	0.063	0.053	1.185	0.362	1
			PW-FS	0.004	0.053	0.079	0.024	1
			NW-FS	-0.033	0.053	-0.632	-0.193	1
		PW-NW	PW-FS	-0.058	0.053	-1.106	-0.338	1
			NW-FS	-0.096	0.053	-1.817	-0.556	1
		PW-FS	NW-FS	-0.037	0.053	-0.711	-0.217	1
Braille	Grp*Cond.	experts, RW-PW	novices, RW-PW	0.092	0.079	1.161	0.636	1
			experts, RW-NW	-0.2	0.087	-2.304	-1.387	1
			novices, RW-NW	0.108	0.079	1.372	0.751	1
			experts, RW-FS	-0.283	0.087	-3.264	-1.965	0.116
			novices, RW-FS	0.083	0.079	1.055	0.578	1
			experts, PW-NW	-0.233	0.087	-2.688	-1.618	0.602
			novices, PW-NW	0.167	0.079	2.11	1.156	1
			experts, PW-FS	-0.317	0.087	-3.648	-2.196	0.035
			novices, PW-FS	0.05	0.079	0.633	0.347	1
			experts, NW-FS	-0.3	0.087	-3.456	-2.081	0.064
			novices, NW-FS	-0.042	0.079	-0.528	-0.289	1
		novices, RW-PW	experts, RW-NW	-0.292	0.079	-3.693	-2.023	0.028
			novices, RW-NW	0.017	0.061	0.272	0.116	1
			experts, RW-FS	-0.375	0.079	-4.748	-2.601	< .001
			novices, RW-FS	-0.008	0.061	-0.136	-0.058	1
			experts, PW-NW	-0.325	0.079	-4.115	-2.254	0.007
			novices, PW-NW	0.075	0.061	1.222	0.52	1
			experts, PW-FS	-0.408	0.079	-5.171	-2.832	< .001
			novices, PW-FS	-0.042	0.061	-0.679	-0.289	1
			experts, NW-FS	-0.392	0.079	-4.96	-2.716	< .001

			novices, NW-FS	-0.133	0.061	-2.172	-0.925	1
		experts, RW-NW	novices, RW-NW	0.308	0.079	3.904	2.138	0.014
			experts, RW-FS	-0.083	0.087	-0.96	-0.578	1
			novices, RW-FS	0.283	0.079	3.588	1.965	0.039
			experts, PW-NW	-0.033	0.087	-0.384	-0.231	1
			novices, PW-NW	0.367	0.079	4.643	2.543	< .001
			experts, PW-FS	-0.117	0.087	-1.344	-0.809	1
			novices, PW-FS	0.25	0.079	3.166	1.734	0.148
			experts, NW-FS	-0.1	0.087	-1.152	-0.694	1
			novices, NW-FS	0.158	0.079	2.005	1.098	1
		novices, RW-NW	experts, RW-FS	-0.392	0.079	-4.96	-2.716	< .001
			novices, RW-FS	-0.025	0.061	-0.407	-0.173	1
			experts, PW-NW	-0.342	0.079	-4.326	-2.37	0.003
			novices, PW-NW	0.058	0.061	0.95	0.405	1
			experts, PW-FS	-0.425	0.079	-5.382	-2.948	< .001
			novices, PW-FS	-0.058	0.061	-0.95	-0.405	1
			experts, NW-FS	-0.408	0.079	-5.171	-2.832	< .001
			novices, NW-FS	-0.15	0.061	-2.444	-1.04	1
		experts, RW-FS	novices, RW-FS	0.367	0.079	4.643	2.543	< .001
			experts, PW-NW	0.05	0.087	0.576	0.347	1
			novices, PW-NW	0.45	0.079	5.698	3.121	< .001
			experts, PW-FS	-0.033	0.087	-0.384	-0.231	1
			novices, PW-FS	0.333	0.079	4.221	2.312	0.004
			experts, NW-FS	-0.017	0.087	-0.192	-0.116	1
			novices, NW-FS	0.242	0.079	3.06	1.676	0.203
		novices, RW-FS	experts, PW-NW	-0.317	0.079	-4.01	-2.196	0.009
			novices, PW-NW	0.083	0.061	1.358	0.578	1
			experts, PW-FS	-0.4	0.079	-5.065	-2.774	< .001
			novices, PW-FS	-0.033	0.061	-0.543	-0.231	1
			experts, NW-FS	-0.383	0.079	-4.854	-2.659	< .001
			novices, NW-FS	-0.125	0.061	-2.036	-0.867	1
		experts, PW-NW	novices, PW-NW	0.4	0.079	5.065	2.774	< .001
			experts, PW-FS	-0.083	0.087	-0.96	-0.578	1
			novices, PW-FS	0.283	0.079	3.588	1.965	0.039
			experts, NW-FS	-0.067	0.087	-0.768	-0.462	1
			novices, NW-FS	0.192	0.079	2.427	1.329	1
		novices, PW-NW	experts, PW-FS	-0.483	0.079	-6.12	-3.352	< .001
			novices, PW-FS	-0.117	0.061	-1.901	-0.809	1
			experts, NW-FS	-0.467	0.079	-5.909	-3.237	< .001
			novices, NW-FS	-0.208	0.061	-3.394	-1.445	0.078
		experts, PW-FS	novices, PW-FS	0.367	0.079	4.643	2.543	< .001
			experts, NW-FS	0.017	0.087	0.192	0.116	1
			novices, NW-FS	0.275	0.079	3.482	1.907	0.055
		novices, PW-FS	experts, NW-FS	-0.35	0.079	-4.432	-2.427	0.002
			novices, NW-FS	-0.092	0.061	-1.493	-0.636	1
		experts, NW-FS	novices, NW-FS	0.258	0.079	3.271	1.792	0.107

Figure 6-2 – Post-hoc t-tests in l-PosTemp. For each effect in the repeated measures ANOVA (rmANOVA), we contrasted the different comparisons to identify the causes of the interaction effects. In Latin script, results are averaged over the levels of 'group'. P-values are adjusted for comparing a family of 15. In Braille script, p-values are adjusted for comparing a family of 66.

Chapter 4. Computational models of vision do not explain expert neural processing of visual Braille

Abstract

The human visual stream adapts to process letters and words at different processing stages (Vinckier et al., 2007), even when the stimuli do not share canonical script features, like Braille (Cerpelloni et al., 2024). These neural results support an interactive account of the Visual Word Form Area (VWFA), where the activations for written stimuli are influenced by interactions with the language network. Here we expand these findings to test the organization of peculiar visual features in computational models. In a first experiment, we looked at how Braille letters are represented in an illiterate Convolutional Neural Network (AlexNet) and compared them to Latin alphabet and to a custom line-based script. We observed a predisposition of the network, pre-trained to perform object recognition, for line-based scripts. Such early result confirms an initial advantage of line junctions over Braille (Bola et al. 2017; Cerpelloni, Collignon, Op de Beeck, 2024). In a second experiment, we trained two benchmark neural network architectures (AlexNet, CORnet Z) to classify words in the Latin script (literacy acquisition) and then in the Braille script (expertise acquisition). We modelled the processing of reading visual Braille and explored the networks representations at different layers. We observed similar degrees of clustering that are explained by the visual properties of the

scripts and not by the network's expertise. Moreover, the representations of linguistic categories do not converge between networks and humans, and neither across network architectures. Overall, the lack of alignment between the visual processing of the computational models and the effect of expertise highlighted by neural data suggests that the fundamental processing of reading cannot be fully explained by the visual characteristics of the script, but necessarily relies on other mechanisms, among which language connections.

Introduction

In humans, the Visual Word Form Area (VWFA) has been found to show a preference for written scripts over other categories typically processed in the ventral occipito-temporal cortex (VOTC) (Cohen et al., 2002; Li et al., 2024). The first models of how VWFA acquires its selectivity for orthography emphasised the progressive integration of line junctions in the visual stream (Dehaene et al., 2005; Vinckier et al., 2007) and found that line vertices posed an advantage both behaviourally and in the brain in both object and word recognition (Szwed et al., 2009, 2011). This led to the hypothesis that line junctions might be necessary to achieve fluent reading (Bola et al., 2017). However, our behavioural findings show that, once script novelty and training paradigm are equated, the advantage of line junctions in learning a new script is limited to the first stages of learning and dampened in its extent (Cerpelloni, Collignon, Op de Beeck, 2024). In parallel, a different hypothesis stressed the involvement of VWFA in other word related tasks (Price & Devlin, 2003) and its interactions and connectivity with areas of the language network (Price &

Devlin, 2003, 2011; Saygin et al., 2016; S. Wang et al., 2022; X. Wang et al., 2018). Recently, high-resolution 7T fMRI has found that VWFA represents different scripts with multiple sub-patches (Zhan et al., 2023), and our previous findings on the neural organization of different scripts (Latin-based and Braille) showed similar coding principles based on the statistical regularities of the stimuli, but relying on a segregated organization of scripts in VWFA and across the visual stream (Cerpelloni et al., 2024).

One approach to separate the contributions of visual and linguistic properties comes from computational models of the visual system. The recent development of Deep Neural Networks (DNNs) made available synthetic models of the visual processing hierarchy, isolated from other cognitive functions that may contribute to object recognition (e.g. top-down connections with language areas). The corpora of studies that apply computational models to reading is still on the rise. Janini and colleagues tested a benchmark Convolutional Neural Network (AlexNet; Krizhevsky et al., 2017) on letter recognition. In particular, they observed that if the network is trained on naturalistic images (ImageNet dataset), the representations of letter identity are better correlated with human behavioural data (reaction time in identifying different letters) than when the network is trained specifically on letters (Janini et al., 2022). Such result supports the idea that the visual hierarchy of object recognition is co-opted by reading (Dehaene et al., 2005). However, this study follows a long tradition of experiments that mostly involve line-based scripts and does not include peculiar case like Braille, sign language, or lip reading. Hannagan, Agrawal, and Dehaene (Agrawal & Dehaene, 2024; Hannagan et al., 2021), in a series of experiments, employed

a network architecture considered more biologically plausible (CORnet Z; Kubilius et al., 2018) to build a synthetic model of VWFA. In the Latin alphabet (French words), the authors observed ordinal and position coding of letters in a word, similar to human findings. This model was tested on different scripts alone and in combinations, providing a solid framework to study reading acquisition and representations in deep neural networks.

In our study, we examined the representations of two deep neural networks when presented with visual Braille stimuli. With this approach, we aimed to isolate the role of visual features (line junctions) from linguistic ones (phonology, semantic meaning) in the task of word recognition. The presentation of Braille letters and words to models can test whether the visual processing hierarchy (without linguistic information) is sufficient to explain the neural representations found in expert human participants. If a network trained to associate Braille and Latin words to the same output class (i.e. same meaning) were to represent those scripts similarly to how expert humans do, the results found for visual Braile reading could then be explained in terms of visual integration. On the other hand, a divergence of results would highlight the role of the language network and its connections with the visual system as a missing factor to achieve human representations. We studied this hypothesis in different network architectures and trainings. In a first experiment, we explored how an illiterate AlexNet (i.e. not trained on alphabetic stimuli) represents letters coming from the Latin alphabet, from Braille, and from a custom script with Braille dots joined into lines (Line Braille). We observed higher similarities between Line Braille and the Latin alphabet, both showing low similarity with Braille. In a second experiment, we then looked at the

representations of full words, whether known or not by the network. We trained AlexNet and CORnet Z models to acquire literacy (if needed), and then introduced Braille or Line Braille alongside the Latin alphabet. Both networks architectures showed an advantage in training a line-based script over Braille. At the end of training (expertise acquisition), we tested the models on the classification of a test set of stimuli in Latin and Braille alphabets where we modulated the linguistic properties of the stimuli (Real Words, Pseudo Words, Non Words, Fake Script). Overall, network expertise determined an increase in dissimilarity across layers, but impacted the clustering of representations (i.e. the difference between two categories over the difference within them) only marginally, and did so differently between network architectures. The networks representations of scripts showed within-script correlations and no clear effects of expertise. Although preliminary, these findings hint at the visual processing of computational models being not sufficient to explain the neural results, highlighting instead that vision alone cannot explain how linguistic stimuli are represented in the brain and arguing in favour of the need of language network interactions during reading.

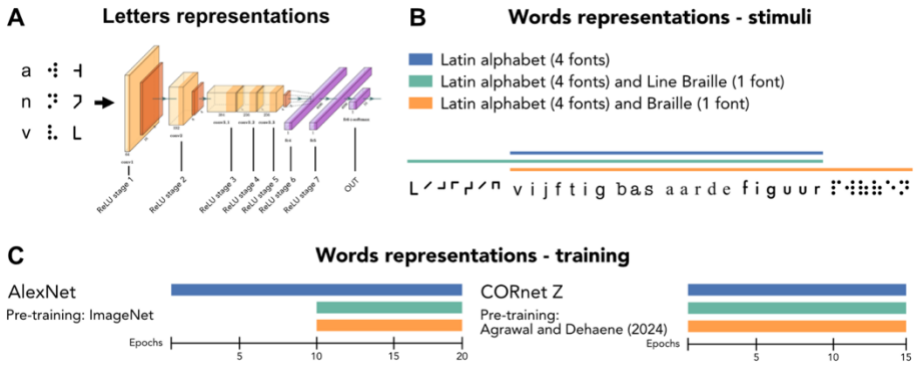

Figure 11 – Design of the experiments. (A) Experiment 1, representations of letters. Letters in Latin alphabet (Arial font), Braille, Line Braille were presented to a AlexNet instance pre-trained on ImageNet. We extracted the representations of each letter at different processing stages. (B,C) Experiment 2, word stimuli and training. (B) Stimuli were presented to the networks into different datasets: Latin alphabet alone, Latin alphabet + Line Braille, Latin alphabet + Braille and examples of stimuli from different fonts (left-to-right: Line Braille, Arial, American Typewriter, Times New Roman, Futura, Braille). (C) Training paradigms of the networks. AlexNet, previously trained only on ImageNet, first underwent literacy acquisition (epochs 1 to 10) and then, in different instances, further trained on one of the datasets (expertise acquisition, epochs 11 to 20). CORnet Z, previous trained by Agrawal and Dehaene on Latin alphabet stimuli (French language), was directly trained on either dataset for the expertise acquisition (epochs 1 to 15).

Methods

Stimuli

To explore the organization of scripts with and without line junctions, we compared letters in the Braille script, in a Latin-based script, and in a custom script made of lines (Line Braille, cf. Chapter 2). We previously used these scripts in a behavioural training experiment (Cerpelloni, Collignon, Op de Beeck, 2024) and showed an initial advantage of lines over dots, but also that such advantage was compensated by the amount of training and the learning of full words. For the Latin alphabet, we used Arial font letter images of the

same size as the custom script. To increase the variability of the dataset, we introduced variations of the thickness and size of the letters presented in the three scripts. Thickness variations are made by changing the area of the dots and lines, adding or removing a pixel outline to the characters at each step. We opted for five levels, consisting of -6, -3, 0 (original image), +3, +6 pixels layers. Similarly, for the size variations we opted to vary it on five levels, creating letters that were -30%, -15%, 0%, +15%, +30% the size of the original letters. All images were 500x500 pixels, with the letters presented in black at the centre of a white background.

To address the contributions of the visual features to reading, we referred to the methods used by Agrawal and Dehaene (2024) and constructed datasets to train and test the network architectures. To train the networks on the acquisition of literacy and expertise, we built three datasets representing the literacy and expertise of the different human populations previously tested. We implemented (1) a condition of Latin-based script alone (LT) to reproduce the processing happening in the naïve participants whose VWFA organization was investigated and in the participants of the behavioural training prior to beginning to learn a new script; (2) a dataset of Latin-based and Braille alphabets (LTBR) to replicate both the group of novice individuals who underwent online learning of visual Braille, and the visual Braille expert participants tested on their neural organization; (3) a dataset of Latin-based and Line Braille alphabets (LTLN), to parallel the group of novice participants who learnt Line Braille in the online training. We chose to maintain the Latin-based alphabets in all the datasets to reflect the training performed by human participants, who learnt the new script without losing exposure to Latin-based

Dutch (and its disproportion). All datasets contain the same one thousand Dutch words from the Dutch Lexicon Project (Brysbaert et al., 2016), but varied in the number of fonts they contained. All words selected were frequent (top 50% of the dataset) and balanced in length, component (i.e. piece of sentence), orthographic neighbours. A dataset comprised of either 1100 or 1375 stimuli per word, depending on the training set: the LT dataset contained only Latin-based alphabet fonts (Arial, Times New Roman, American typewriter, Futura), while the LTBR and the LTLN datasets added a font of Braille or Line Braille to the four fonts of the LT dataset. Within a font, each word was varied in size (-30%, -15%, 0%, +15%, +30% variations from the original size), and shifted on the x- (-50, -40, -30, -20, -10, 0, 10, 20, 30, 40, 50 pixels) and y-axes (-40, -20, 0, 20, 40 pixels) from the center of the image. To test the within-layer representations in a network, we constructed a test set based on the stimuli from our fMRI experiment on the organization of linguistic categories in VWFA (Cerpelloni et al., 2024). This test set included twelve stimuli from each linguistic category: (1) Real Words (RW) from the training set classes; (2) Pseudo Words (PW) built using unipseudo (New et al., 2024) from a subset of the training set; (3) Non Words (NW), strings of consonants; (4) Fake Script stimuli (FS). While we adapted the Real Words and Pseudo Words sets (related to each other) to Dutch and to the specific requirements of this study, we used the same stimuli of the fMRI experiment for the sets of Non Words and Fake Script. All stimuli were presented in both Latin-based (Arial) and Braille alphabets, and underwent the same size variations and axes shifts of the training sets.

Deep Neural Network architectures and training

We implemented two neural network architectures from previous studies, AlexNet (Krizhevsky et al., 2017) and CORnet Z (Kubilius et al., 2018), and performed in-house training on the datasets described above using Pytorch libraries and available weights.

AlexNet

In experiment 1, we implemented AlexNet for its wide use in computational neuroscience, which allows for easy comparisons with other studies on object recognition, in particular the study of letter processing of Janini and colleagues (2022) on the classification of letters by an ImageNet-trained AlexNet. Building on these results, we investigated the early representations of letters by presenting single-letter stimuli in three scripts (Latin-based, Braille, Line Braille) to an AlexNet architecture pre-trained on ImageNet. In experiment 2, to investigate the processing of full words in canonical and peculiar scripts, we performed two trainings. We implemented instances (N = 5) of the network pre-trained on ImageNet (Russakovsky et al., 2015), referred to as “illiterate”, and reset the output classes to extract classification performance at the end of the training. Since AlexNet models were not previously trained on words but only on ImageNet, we first trained our models on the Latin-alone training set (literacy acquisition). We trained each instance on a total of 20 epochs (learning rate = 1e-4; batch size = 100; training set = 70% of the full dataset). After 10 epochs, we reached a classification accuracy above 95% (literacy acquisition) and we introduced expertise for a novel script. The same network instance was copied three times and each copy was further trained, with the same parameters, for another 10 epochs on one of the three datasets,

introducing Braille (further trained on the Lain-based and Braille alphabets; N = 5), Line Braille (further trained on the Lain-based and Braille alphabets; N = 5), or continuing with the Latin alphabet training (further trained on the Latin-alone dataset; N = 5). Although in this last case the network training performance reached a plateau after 10 epochs and did not increase with additional training, we continued training to equate the amount of training with the expert ones. Literacy and expertise acquisition resulted in five instances of each network that mimic the results of the human studies.

CORnet Z

CORnet Z was chosen to replicate as closely as possible the works of Hannagan, Agrawal, and Dehaene (Agrawal & Dehaene, 2024; Hannagan et al., 2021) focused on building a synthetic model of VWFA. In experiment 2, given the availability of the data from these studies, we implemented the French-literate network presented in Agrawal and Dehaene (2024) as a starting model for our training, adapting the number of output units (two thousand in their study, fitting both words and ImageNet classes) to our one-thousand training classes. We opted for one network, and did not perform a comparison of the English-literate, or the bilingual French-English networks the authors implemented, as the Latin-alphabet letters are shared among scripts. CORnet models, already trained on a dataset of words and therefore considered literate, were not trained in two steps (literacy and then expertise) but only performed one training step. Each network instance, with literate weights coming from one of the five instances of Agrawal and Dehaene's study, was copied three times and each copy was trained on one of the three words datasets. We further trained on the Latin alphabet alone to obtain a benchmark of the classification

accuracy for the new classes of Dutch words introduced. We trained each instance of CORnet for 15 epochs, using the same parameters used for AlexNet training (learning rate = 1e-4; batch size = 100; training set = 70% of the full dataset), to achieve similar levels of accuracy across trainings and networks. We extended the training epochs in CORnet, compared to AlexNet, as the condition of training Latin alphabet and Braille showed lower accuracy at epoch 10. The additional epochs served to reach comparable levels of accuracy between network trainings.

Testing

After training, whether on Latin-based alphabet alone (naïve networks) or introducing a novel script (Braille expert), we tested the networks against our linguistic conditions dataset of Latin-based and Braille scripts. For both AlexNet and CORnet networks, we measured the features activations from the nodes used by previous studies (AlexNet: ReLU stages of each layer, and output; CORnet: output layer of each CORBlock, linear and output layers of decoder) for each stimulus image presented to the network.

Statistical analyses

In experiment 1, to measure the dissimilarity between letters identities, we referred to the methods that Janini and colleagues used in their second experiment, investigating the representations of letter identity. We computed the Euclidean distance between stimuli, comparing one stimulus averaged across variations to all the variations of another stimulus, repeating the same in the opposite direction to then average the two distances obtained (Janini et al., 2022), resulting in a representational dissimilarity matrix (RDM). In

experiment 2, we used the same method on the activations for stimuli of different linguistic categories. We computed a measure of the clustering from the RDMs as the average dissimilarity in comparisons between conditions minus the dissimilarity within conditions, divided by the overall average dissimilarity of representations in the four different linguistic conditions (Real Words, Pseudo Words, Non Words, Fake Script). We computed this measure of clustering at each layer of the networks employed. Lastly, we averaged the clustering measures within a layer, to obtain a single value representing the clustering at a given stage. To probe for different processing based on the expertise, we performed a repeated measures ANOVA on the clustering measure across layers, scripts, and networks' expertise. To compare our results to previous studies, we additionally measured the mean dissimilarity (1-correlation) across activations for each stimulus at each layer, following the methods of Agrawal and Dehaene (2024). The single value per layer indicates the overall enhancement in neural responses comparing script and expertise conditions. To explore the relationships between stimuli representations of the different scripts and expertises, we correlated the dissimilarity matrices of the categories at the output layer within a network architecture. Given the small sample of network instances, we followed the method used previously (Cerpelloni et al., 2024) and computed non-parametric statistics by performing bootstrapping to obtain a null distribution of the correlations between matrices.

Results

Letter processing depends on the presence of line junctions

We presented letters in the Latin alphabet (Arial font), in Braille, and in a custom Line Braille script to an instance of AlexNet pre-trained on ImageNet but illiterate (i.e. not directly trained on written material) and measured the network activations at different processing stages (Figure 2). We observed striking similarities between the Latin alphabet and Line Braille, two scripts based on line junctions, with the scripts clustering together, while Braille shows progressively higher dissimilarity compared to those two scripts. This pattern of dissimilarities between dots and lines is present across all layers, and becomes more marked from the middle layers. This observation hints at a preferential processing of line junctions in a network trained solely on natural objects, and at different visual representations for a script that does not share those basic features.

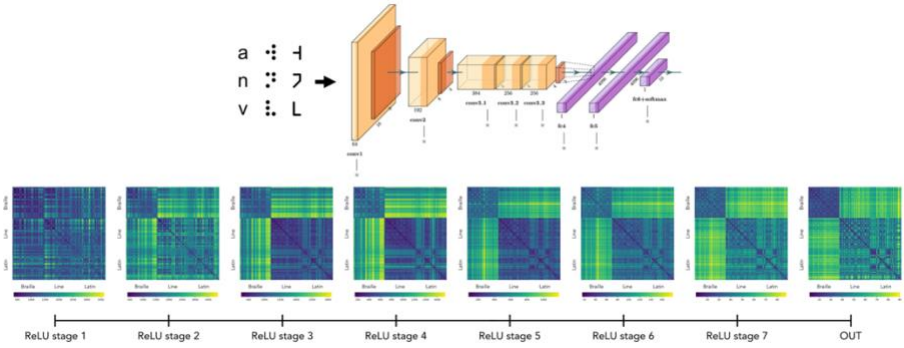

Figure 2 – Illiterate AlexNet representations of letters in different scripts. Letters in Braille, in Line Braille, in Latin alphabet presented to the network. Activations were extracted given stages (ReLU stages and output) for all the individual letters in their size and thickness variations. We compared the representations for the letter identity (across variations) across scripts by computing the Euclidean distance between stimuli. Bottom row shows the Representational Dissimilarity Matrices (RDMs) at different processing stages for all the letters of Braille, Line Braille, Latin alphabet respectively. From early layers, a marked difference between Braille and line-based scripts emerges and remains at the output layer.

Expertise acquisition shows benefits of line-based scripts over Braille

We trained AlexNet (Figure 3A) to first classify Dutch words presented in four Latin-based fonts (literacy acquisition) and then, in some cases, introduced either Braille or Line Braille (joining the dots to form line junctions) fonts alongside the Latin alphabet fonts, mapping the new script on the same output units (expertise acquisition). We observed a drop of performance in both expertise conditions, probably due to task-switching costs, stronger for Braille (average drop: -0.4056) than for the line-based script (average drop: -0.1366). In the case of CORnet Z training (Figure 3B), starting from literate weights, we trained in parallel different instances on the acquisition of literacy (learning words in the Latin alphabet) and on the acquisition of expertise for a novel script (adding Braille or Line Braille to the Latin alphabet words). We observed, compared to the training on the Dutch words presented in four Latin-based fonts alone, a temporal delay in the classification accuracy for Line Braille expertise, which is compensated at epoch 10 (when there is a plateau of the Latin alphabet dataset). The Braille expertise acquisition, however, shows a more marked delay, with an average classification accuracy at epoch 10 of 0.5978, much lower than the line-based trained networks (Latin alphabet alone: 0.9689; Line Braille expertise: 0.9435). Overall, although with differences in the training paradigms across networks, both AlexNet and CORnet architectures present an advantage in introducing a scripts that presents the canonical line junctions rather than a script without them. This result can be compared with our behavioural findings in the training of human participants, where we found a limited advantage of learning Line Braille over

dot-based Braille in the first session of training (Cerpelloni, Collignon, Op de Beeck, 2024), in this case accentuated by the purely visual training.

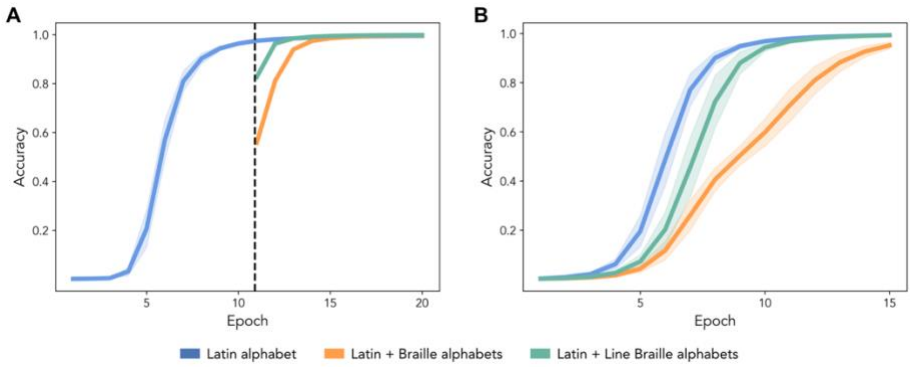

Figure 3 – training regimes for the network architectures. (A) AlexNet architecture. Five instances of AlexNet were trained for 20 epochs to classify a dataset of words in the Latin alphabet. After 10 epochs, we introduced either Braille or Line Braille to the dataset of Latin alphabet words, training on both scripts. After an initial drop in accuracy, we observe accuracy to reach ceiling within 5 epochs for both scripts. (B) CORnet Z architecture. We started from pre-trained weights from Agrawal and Dehaene (2024), where the authors trained the architecture to become literate. We further trained five instances of the network on each dataset of our experiment. Introducing Line Braille does not disrupt learning compared to the Latin alphabet dataset, while the addition of Braille results in a lower accuracy throughout the training epochs. In both figures, the shaded area indicates 95% confidence interval.

Expertise acquisition does not influence word representations

We replicated the analyses performed by Agrawal and Dehaene (2024) and computed the mean dissimilarity of representations for each script at each processing stage (Figures 4A and 4B). Both networks and alphabets show an increase in the dissimilarity across layers, even when a naïve network processes Braille stimuli. However, there is a difference between trained and not trained scripts in AlexNet: Latin alphabet stimuli, trained in both network implementations, and Braille stimuli in the expert network showed higher mean dissimilarity compared to the naïve network activations for Braille. In

CORnet models we observed an increase in the mean dissimilarity for the Latin alphabet from the layer associated to IT, but delayed increase in the case of an expert network processing Braille, which happens only in the last stage. This observation hints at the possibility of improvements in the representations at higher-level processing stages.

We then computed, for each network and layer, a measure of clustering (average dissimilarity in comparisons between conditions minus the dissimilarity within conditions, divided by the overall average dissimilarity) based on the dissimilarity of the activations in response to the test stimulus images. Using clustering, we performed a repeated measures ANOVA (2 scripts – Braille and Latin-based * 2 network trainings – expert and naïve to Braille * 6 or 8 layers, depending on the network). Overall, the statistical analyses show very similar results (Figures 4C and 4D), but with different underlying meanings. More specifically, in AlexNet models (Figure 4C), there is a general increase in the clustering of representations associated to the increase in the complexity of the processing stage (main effect of layer: $F_{(7,56)} = 1626$, $p < 0.001$), and a higher clustering for the Latin-based alphabet than for Braille (main effect of script: $F_{(1,8)} = 3848$, $p < 0.001$). However, whether the network was explicitly trained to associate Braille words to Latin-based categories does not play a role in how the network represents the different scripts (main effect of network's expertise: $F_{(1,8)} = 3.05$, $p = 0.12$). Significant interaction effects are present between expertise and all the other variables (expertise * layer: $F_{(7,56)} = 30.53$, $p < 0.001$; expertise * script: $F_{(1,8)} = 35.97$, $p < 0.001$; layer * script: $F_{(7,56)} = 408.77$, $p < 0.001$; expertise * layer * script: $F_{(7,56)} = 109.36$, $p < 0.001$). In particular, the interactions with expertise seem to be driven by the increase in the clustering of Braille stimuli happening in the fully-connected layers

(ReLU stage 6 and 7, and output) of naïve AlexNet. This increase can be attributed to the clustering measure itself: while the score is high, it underlies overall low dissimilarities between the stimuli of one category, meaning that the categories are not processed differently from each other. Similarly, the clustering of representations in CORnet Z models (Figure 4D) is mainly determined by the processing stage (main effect of layer: $F_{(5,40)} = 954.95$, $p < 0.001$) and by the script processed (main effect of script: $F_{(1,8)} = 798.82$, $p < 0.001$), but not by the network expertise with Braille (main effect of expertise: $F_{(1,8)} = 1.66$, $p = 0.23$). Moreover, significant interactions between expertise, layer and script (expertise * layer: $F_{(5,40)} = 41.72$, $p < 0.001$; expertise * script: $F_{(1,8)} = 211.47$, $p < 0.001$; layer * script: $F_{(5,40)} = 39.22$, $p < 0.001$; expertise * layer *script: 16.83, $p < 0.001$) are present. Of particular importance, the interaction between expertise and script differs from what we observed in AlexNet. In this case, the interaction is driven by the increase in clustering of Braille representations that happens in the Braille expert network from the IT layer, an increase that is not confounded by the overall category dissimilarity and that may correspond to what we observed in human participants in VWFA. In both networks, the output representations of the linguistic categories do not show clear correlation patterns as shown in human results (Figures 4E and 4F). AlexNet (Figure 4E) presents significant correlations within scripts (expert, Latin – naïve, Latin: $r = 0.95$, $p_{FDR} < 0.001$; expert, Braille – naïve, Braille: $r = 0.79$, $p_{FDR} < 0.001$), and only between Latin alphabet in the expert network and Braille alphabet in the naïve network (expert, Latin – naïve, Braille: $r = 0.50$, $p_{FDR} = 0.0002$). All the other correlations within and between networks (all $p_{FDR} > 0.08$) are not significant. A similar trend appears in CORnet models, with significant within-script correlations (expert, Latin – naïve, Latin: $r = 0.99$, p_{FDR}

< 0.001 ; expert, Braille – naïve, Braille: r = 0.89, p_{FDR} < 0.001) and with a significant correlation between the expert network's organization of Braille and the naïve network's organization of the Latin alphabet stimuli (expert, Braille – naïve, Latin: r = 0.39, p_{FDR} = 0.004). All other correlation are not significant (all p_{FDR} > 0.09). Overall, different architectures show different results but with a common trend of mostly showing correlations within a script. Both networks converge in showing no main effects of expertise, diverging in what drives the significant interactions with other factors.

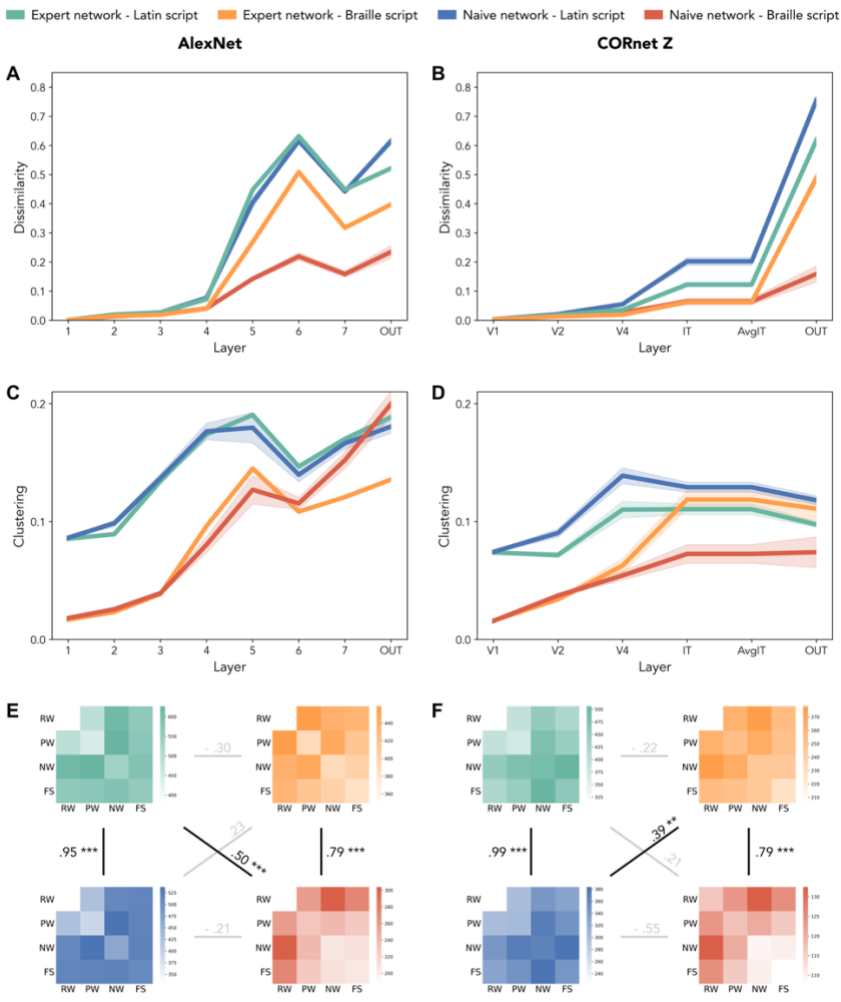

Figure 4 – Network representations of words. (A, B) Mean dissimilarity across stimuli. AlexNet (A) and CORnet (B) both show an increase in the mean dissimilarity between test stimuli of a script in the middle to last layers, similarly to what shown in previous work. This increase is more marked for scripts that have been presented during training compared to novel scripts. (C, D) Clustering of test categories across layers. AlexNet (C) shows higher clustering (more dissimilarity in the off-diagonal compared to the diagonal) for the Braille script when not trained on it, while CORnet (D) shows an effect of expertise for Braille stimuli, with comparable levels of clustering to the Latin alphabet (extensively trained in all conditions). In A, B, C, D, the shaded area indicates 95% confidence interval. (E, F) Representational Dissimilarity matrices (RDMs) of the dissimilarities averaged across categories, for the two scripts in the two conditions at the output layer, for AlexNet (E) and CORnet (F). No pattern emerges from either network architecture. Non-significant correlations are shown in grey, significant correlations are shown in black (* = $p_{FDR} < 0.05$; ** = $p_{FDR} < 0.01$; *** = $p_{FDR} < 0.001$).

Discussion

We explored the visual processing of Braille letters and words using computational models previously applied to the same classes of stimuli. We first presented an instance of AlexNet, pre-trained on ImageNet, with letters in the Latin alphabet (Arial font), in Braille, and in a custom script based on line junctions (Line Braille). We observed that the absence of line junctions, as it is the case in Braille, led to higher dissimilarity of the letters representations, and that the presence of line junctions in Line Braille led to closer representations to Latin alphabet, also composed of line junctions, rather than Braille, a script sharing the same overall structure. We then trained AlexNet to process Latin alphabet words, later adding Braille words, and literate CORnet Z models to process either Latin alphabet words or Latin and Braille alphabet words. We observed systematic delays in the learning of Braille compared to learning Line Braille. We tested the networks naïve and expert to Braille on stimuli with different linguistic statistics in both Latin alphabet and Braille and found incongruent results on how the representations of these stimuli are organized, but a common lack of main expertise effects in the progression of clustering and in the representations at the output layers. CORnet Z showed a significant interaction between expertise and script that can hint at an expert processing of Braille in the later stages (i.e. IT). In both networks, the script representations mainly correlated within scripts and did not highlight the trend due to the expertise that was found in humans.

Letters processing in AlexNet showed marked differences between Braille, a script based on dots, and line-based scripts, not only Latin alphabet but also

Line Braille. This early result is particularly interesting for two reasons related to the way these scripts are made. Line Braille is derived from Braille and consists of the same structure expressed through line junctions instead of the canonical dot arrangements. Thus, given the nature of these two scripts, one would expect them to share a high degree of similarity. Instead, Line Braille appears to be represented more similarly to Latin alphabet. This result highlights the tuning of the computational visual hierarchy for line junctions (Agrawal & Dehaene, 2024; Hannagan et al., 2021; Janini et al., 2022), and also complements our previous behavioural findings about Braille learning compared to Line Braille. In this context, the initial advantage of Line Braille may be explained by the perceptual differences of the scripts and by the high similarity of Line Braille to the Latin alphabet. Moreover, the limited similarity between Braille and Line Braille additionally shows the lack of integration of dots into lines that is thought to impair visual reading of Braille. More in general, the dissimilarity between Braille and line-based letters shows a gap between the deep neural network representations, based solely on visual processing, and the human neural representations, where visual areas are connected to linguistic processing areas.

Studies on computational models applied to the processing of written words are limited. Agrawal and Dehaene (2024) tested the expertise of a network, and the effect it has on the script-specific representations, by comparing the activations of their models trained on either Telugu or Malayalam and testing them on bigrams from both scripts. The authors observed that the mean dissimilarity increased from the layer associated to area V4 only when the script matched the network expertise. We observed a similar increase starting

from layer IT of the CORnet models, paralleled by the increase in mean dissimilarity from the middle layers of AlexNet for all the scripts trained, Latin alphabet in the expert and naïve networks, Braille in the expert network. Interestingly, the increased dissimilarity, combined with the different visual representations of Braille letters, may already point to the output category as a driving factor for the internal representations, mimicking a lexical entry in the semantic representations in the brain. We found clustering to be mainly determined by the script processed and only marginally by the expertise of the network. Nonetheless, Braille-expert CORnet models show similar clustering for Braille and for Latin alphabet stimuli, indicating tuning to the statistical redundancies, with or without lines. This ability of the expert network can be interpreted in two alternative manners. On one hand, the visual processing of Braille by a network may indicate that the visual properties of a stimulus, regardless the presence or absence of line junctions, are sufficiently informative. On the other hand, the output class onto which a Braille word is mapped can be seen as a lexical entry, and the output layer as semantic space. To disentangle the issues, we can look at the representations at the output layer. The representations of different scripts for the different expertise do not align with the patterns of correlations found in neural data. Future analyses on the matter should look more in depth at the representations and the different categories of linguistic stimuli, comparing them with a model of linguistic distance and applying the methods of the multivariate fMRI studies to study the cross-script relationships between Latin alphabet and Braille, and relating them to the ones for Line Braille.

Overall, these preliminary results indicate that visual strategies to process Braille letters in artificial networks do not explain the results found in human behaviour when participants learnt Braille (Bola et al., 2017; Cerpelloni, Collignon, Op de Beeck, 2024). Additionally, Braille words processing hints at a clustering of stimuli with different statistical regularities (linguistic properties) that limitedly depends on whether or not a visual network was trained on Braille and instead mainly relies on the visual features of the stimuli. The different relevance of visual and linguistic properties may lead to differences compared to the neural encoding as a result of expertise, as we previously found in multivariate fMRI analyses (Cerpelloni et al., 2024). This discrepancy hints at a possible role of linguistic processing in the visual brain, given that such processing is not included in visual artificial networks. Further simulation studies will pinpoint the necessary and sufficient conditions to reproduce the effect of expertise in the human brain.

References

Agrawal, A., & Dehaene, S. (2024). Cracking the neural code for word recognition in convolutional neural networks. *PLOS Computational Biology*, *20*(9), e1012430. https://doi.org/10.1371/journal.pcbi.1012430

Bola, Ł., Radziun, D., Siuda-Krzywicka, K., Sowa, J. E., Paplińska, M., Sumera, E., & Szwed, M. (2017). Universal Visual Features Might Be Necessary for Fluent Reading. A Longitudinal Study of Visual Reading in Braille and Cyrillic Alphabets. *Frontiers in Psychology*, *8*. https://doi.org/10.3389/fpsyg.2017.00514

Brysbaert, M., Stevens, M., Mandera, P., & Keuleers, E. (2016). The impact of word prevalence on lexical decision times: Evidence from the Dutch Lexicon Project 2. *Journal of Experimental Psychology: Human Perception and Performance*, *42*(3), 441–458. https://doi.org/10.1037/xhp0000159

Cerpelloni, F., Collignon, O., & Op De Beeck, H. (2024). *Connecting the dots: Similar visual orthographic acquisition for Braille or line junctions*. https://doi.org/10.31234/osf.io/y3sg2

Cerpelloni, F., Van Audenhaege, A., Matuszewski, J., Gau, R., Battal, C., Falagiarda, F., Op de Beeck, H., & Collignon, O. (2024). Expertise in reading visual Braille co-opts the reading network. *bioRxiv*, 2024.05.08.593104. https://doi.org/10.1101/2024.05.08.593104

Cohen, L., Lehéricy, S., Chochon, F., Lemer, C., Rivaud, S., & Dehaene, S. (2002). Language-specific tuning of visual cortex? Functional properties of the

Visual Word Form Area. *Brain*, *125*(5), 1054–1069. https://doi.org/10.1093/brain/awf094

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

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

Janini, D., Hamblin, C., Deza, A., & Konkle, T. (2022). General object-based features account for letter perception. *PLOS Computational Biology*, *18*(9), e1010522. https://doi.org/10.1371/journal.pcbi.1010522

Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2017). ImageNet classification with deep convolutional neural networks. *Communications of the ACM*, *60*(6), 84–90. https://doi.org/10.1145/3065386

Kubilius, J., Schrimpf, M., Nayebi, A., Bear, D., Yamins, D. L. K., & DiCarlo, J. J. (2018). *CORnet: Modeling the Neural Mechanisms of Core Object Recognition*. Neuroscience. https://doi.org/10.1101/408385

Li, J., Hiersche, K. J., & Saygin, Z. M. (2024). Demystifying visual word form area visual and nonvisual response properties with precision fMRI. *iScience*, *27*(12), 111481. https://doi.org/10.1016/j.isci.2024.111481

New, B., Bourgin, J., Barra, J., & Pallier, C. (2024). UniPseudo: A universal pseudoword generator. *Quarterly Journal of Experimental Psychology*, *77*(2), 278–286. https://doi.org/10.1177/17470218231164373

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

Price, C. J., & Devlin, J. T. (2011). The Interactive Account of ventral occipitotemporal contributions to reading. *Trends in Cognitive Sciences*, *15*(6), 246–253. https://doi.org/10.1016/j.tics.2011.04.001

Russakovsky, O., Deng, J., Su, H., Krause, J., Satheesh, S., Ma, S., Huang, Z., Karpathy, A., Khosla, A., Bernstein, M., Berg, A. C., & Fei-Fei, L. (2015). ImageNet Large Scale Visual Recognition Challenge. *International Journal of Computer Vision*, *115*(3), 211–252. https://doi.org/10.1007/s11263-015-0816-y

Saygin, Z. M., Osher, D. E., Norton, E. S., Youssoufian, D. A., Beach, S. D., Feather, J., Gaab, N., Gabrieli, J. D. E., & Kanwisher, N. (2016). Connectivity precedes function in the development of the visual word form area. *Nature Neuroscience*, *19*(9), 1250–1255. https://doi.org/10.1038/nn.4354

Szwed, M., Cohen, L., Qiao, E., & Dehaene, S. (2009). The role of invariant line junctions in object and visual word recognition. *Vision Research*, *49*(7), 718–725. https://doi.org/10.1016/j.visres.2009.01.003

Szwed, M., Dehaene, S., Kleinschmidt, A., Eger, E., Valabrègue, R., Amadon, A., & Cohen, L. (2011). Specialization for written words over objects in the visual cortex. *NeuroImage*, *56*(1), 330–344. https://doi.org/10.1016/j.neuroimage.2011.01.073

Vinckier, F., Dehaene, S., Jobert, A., Dubus, J. P., Sigman, M., & Cohen, L. (2007). Hierarchical Coding of Letter Strings in the Ventral Stream: Dissecting

the Inner Organization of the Visual Word-Form System. *Neuron*, *55*(1), 143–156. https://doi.org/10.1016/j.neuron.2007.05.031

Wang, S., Planton, S., Chanoine, V., Sein, J., Anton, J.-L., Nazarian, B., Dubarry, A.-S., Pallier, C., & Pattamadilok, C. (2022). Graph theoretical analysis reveals the functional role of the left ventral occipito-temporal cortex in speech processing. *Scientific Reports*, *12*(1), 20028. https://doi.org/10.1038/s41598-022-24056-1

Wang, X., Xu, Y., Wang, Y., Zeng, Y., Zhang, J., Ling, Z., & Bi, Y. (2018). Representational similarity analysis reveals task-dependent semantic influence of the visual word form area. *Scientific Reports*, *8*(1), 3047. https://doi.org/10.1038/s41598-018-21062-0

Zhan, M., Pallier, C., Agrawal, A., Dehaene, S., & Cohen, L. (2023). Does the visual word form area split in bilingual readers? A millimeter-scale 7-T fMRI study. *Science Advances*.

Chapter 5. Discussion

In this dissertation, I presented three studies which employed multiple complementary approaches to study the mechanisms underlying visual Braille reading. The absence of line junctions, visual features shared among scripts (Changizi et al., 2006), together with the unaltered linguistic information, makes visual Braille a candidate model to discern the contributions of vision and language and to clarify the two contrasting theories: orthographic tuning hypothesis (Dehaene et al., 2005; Dehaene & Cohen, 2011; McCandliss et al., 2003) and interactive account (Price & Devlin, 2003, 2004, 2011) of VWFA.

Behavioural learning of visual Braille

In chapter 2, we explored the behavioural changes associated with learning new scripts. Two groups of individuals performed an online experiment where they were asked to learn either visual Braille or Line Braille, a custom script where we joined the dots to form line junctions. Regardless of the alphabet, participants performed the same training. Over four sessions in consecutive days, they first learnt to associate letters in the novel alphabet to their native Latin alphabet (Dutch) and later, in successive sessions, viewed full words (training phase). At the end of each session, participants were asked to transcribe words and pseudo-words from the novel alphabet to their native one (test phase). Across training and test phases, a general similarity between the two scripts emerged, with almost no main effects of group differences among the behavioural measures recorded. The main differences between

scripts were minor and consisted of Braille having a slight delay in the accuracy during the test and an increased sensitivity to stimulus length. Previous studies showed the importance of line junctions and vertices in object and word recognition (Szwed et al., 2009, 2011) and in the learning of a new script (Bola et al., 2017). In particular, the result of Bola and colleagues shows a strong advantage in learning a script based on line junctions over visual Braille (Bola et al., 2017). However, in our experiments we levelled the training paradigm and the novelty of the alphabets between the scripts, finding dampened differences between line- and dot-based scripts. Here, the overall lack of group differences in the learning behaviour shows the strength of the linguistic content mapping as the determinant factor in the acquisition of a new script, with a relative modulation of the visual features of the stimuli learnt.

Neural organization of visual Braille

At the neural level, in chapter 3 I presented experiments done with expert visual readers of Braille. This group, together with a larger sample of age-matched controls naïve to Braille, was asked to perform two experiments while their brain activity was recorded through fMRI. First, we localized the brain regions involved in processing scripts. Individuals with many years of experience reading Braille visually showed neural activations for Braille words in a site previously, using independent data, identified as the Visual Word Form Area (VWFA). This preference for Braille words over scrambled dot arrangements appeared consistently among expert readers and was virtually absent in naïve controls. The experts' activation of VWFA in response to

visually presented Braille stimuli is in contrast to the orthographic tuning hypothesis (Dehaene et al., 2005; Dehaene & Cohen, 2011), in which VWFA integrates line junctions, but in support of a more general tuning of VWFA for alphabetic material, irrespective of the stimulus characteristics (Price & Devlin, 2003, 2011). This finding aligns with many studies showing VWFA activation for different alphabets (Baker et al., 2007; Vogel et al., 2012; Xue & Poldrack, 2007), different sensory modalities (Büchel et al., 1998; Pattamadilok et al., 2019; Planton et al., 2019; Reich et al., 2011), different types of stimuli associated to linguistic content (Martin et al., 2019; Moore et al., 2014; Van Audenhaege et al., 2024), and adds a fine comparison of different scripts that share language information and sensory modality but differ significantly in their visual features.

In a second fMRI experiment, participants processed stimuli with different linguistic content. In VWFA, individuals showed a fine-scale multivariate pattern of activity in response to Braille and Latin alphabet stimuli based on their familiar orthographies (in expert readers for both Latin and Braille alphabets, in naïve controls only for the Latin alphabet). Such patterns were strongly correlated between the different scripts and aligned to a model of the linguistic content. However, we did not find significant cross-decoding between script, meaning that the similar processing between the Latin alphabet and Braille found in expert participants did not reveal a common representation of linguistic content, abstracted from the visual features. This segregation of sensory characteristics was also recently observed in lip-reading for phonemes and visemes (Van Audenhaege et al., 2024) and may underlie different neuronal populations coding different inputs that convey the same

information. Moreover, such trend was present in several areas of the reading network. Both primary visual cortex (V1) and lateral occipital complex (LOC), regions involved in object recognition, revealed such organization pattern even in the absence of univariate preference for words. These findings are in line with previous studies showing preferential activation for trained scripts in V1 (Szwed et al., 2014) and script processing in LOC (Haupt et al., 2024). In the language network, the left Posterior Temporal region showed a similar pattern between neural organizations, although not significantly correlated with the linguistic model, and additionally presented a convergence of the representations into a common organization regardless of the visual properties. These results show that the linguistic content determines the coding principles used to organize information in the areas involved in reading, supporting an interactive account of VWFA (Price & Devlin, 2003, 2011). However, in the visual system, these principles are enacted with the modulation of the visual properties. Additionally, the evident pattern emerging from different regions of the reading network shows the strength of Multi-Variate Pattern Analysis (MVPA) as a tool to investigate the neural mechanisms on a finer scale than the overall presence or absence of a selective response. The few studies that applied MVPA to the pattern of responses in VWFA found similar results, in the way information is organized based on linguistic context (X. Wang et al., 2018) and in the way linguistic properties are stored in relation to the stimulus sensorial features (Van Audenhaege et al., 2024).

Computational modelling of visual Braille

Chapter 4 presented computational experiments involving models of the visual system. We tested two Deep Neural Network (DNN) architectures on the recognition of letters and words, to assess the contribution of vision alone to the behavioural and neural results, excluding the potential impact of language knowledge. First, we presented letters to an instance of AlexNet, trained to perform object-recognition but not explicitly on words. The representations of Braille, Line Braille, and Latin alphabet letters showed high similarities between Latin and Line Braille alphabets, with higher dissimilarities between these line-based scripts and Braille. The pattern of similarities in line-based scripts cluster together and distantly from Braille supports the behavioural results (chapter 2) and can explain the differences in transcription accuracy at test, heavily influenced by the different performance at the first session, after training on individual letters.

We then trained two DNN architectures (AlexNet, Krizhevsky et al., 2017; CORnet Z, Kubilius et al., 2018) on datasets that reproduced the experience of the human participants tested in the previous behavioural and neuroimaging experiments. These datasets included words in the Latin alphabet alone (cf. chapter 3, naïve controls) and in combination with either Braille (cf. chapter 2, Braille learners and chapter 3, Braille experts) or Line Braille (cf. chapter 2, Line Braille learners) mapped on the same output classes as the Latin alphabet words. AlexNet models, trained first on the Latin alphabet alone and then on either extended dataset, showed an initial impairment in the learning accuracy for Braille compared to Line Braille. Such drop in performance is associated

with the absence of line junctions in Braille. CORnet models, already available as pre-trained on the Latin alphabet (Agrawal & Dehaene, 2024), were trained in either one of the three datasets. The learning curve of Line Braille showed a slight delay compared to the Latin alphabet alone, while Braille learning was associated to a marked decrease in the accuracy, more substantial than Line Braille. The difference in the learning curves shows, similarly to the results from the letters representations, large processing differences based on the visual properties of these scripts.

Lastly, we tested networks that were previously trained on Braille words (expert networks) and instances trained solely on the Latin alphabet (naïve networks), and presented them with stimuli in both Braille and Latin alphabets that possessed different linguistic content (cf. chapter 3, MVPA stimuli). The mean dissimilarity between conditions, which indicates the sensitivity to the linguistic categories, increased with the processing hierarchy, similarly to what was found in a previous study that developed the CORnet literate network (Agrawal & Dehaene, 2024). However, the dissimilarity increase was higher for the Latin alphabet than for Braille in both networks, regardless their expertise. The multivariate representations of linguistic content did not highlight correlations between familiar scripts or with a model of linguistic processing that would resemble the human results. Instead, they showed similarities based on the visual properties of the scripts and not on the training-induced expertise. The lack of an effect of expertise shows how a synthetic visual system is not sufficient to elucidate the processing of visual Braille as a script and argues instead for the interactions with the language network as a determinant factor to explain human reading.

Connecting the dots reveals the influence of language and vision in reading

The three complementary studies (chapters 2, 3, 4), derivate from one another, can be joined together to complete the picture of visual Braille processing and the contributions of language and vision to reading.

The participants of the behaviour and neuroimaging studies can be viewed as stages on the same temporal continuum, respectively at the beginning of learning and at different timepoints depending on the years of experience. Although from the sample size in the neuroimaging experiments we could not draw meaningful correlations between expertise and neural activations (see chapter 3, discussion), we can identify a trend in the type of properties that determine the participants' results. In both cases, the main drive of the responses arises from the linguistic content of the stimuli and is modulated by the visual features. Indeed, the behavioural data indicates that learning Braille presents minor disadvantages compared to Line Braille, with a limited delay in the accuracy and higher sensitivity to the length of the stimuli. Similarly, the broad multivariate organization of Braille and Latin alphabet shows similar patterns between scripts based on the linguistic knowledge rather than a processing based on visual features, which however determines a separation of representations. In sum, the role of language and the strength of the mapping between novel and consolidated orthographies seems to overcome the visual differences of the orthographies and determine both learning and neural organization, however without excluding the visual aspects completely.

Parallels can also be drawn between human and computational results, which were in fact direct replications of the behavioural and neuroimaging experiments. First, synthetic models can expand and explain the results of letter representations and word learning. In human participants, we observed similar learning curves between scripts in the transcription accuracy during the test phase. However, an interaction emerges between scripts and sessions, due to different learning trajectories for Braille and Line Braille. Post-hoc t-tests identify a significant difference in the first session, when participants are only trained on letters before being tested on full words, but this trend is not confirmed after correction for multiple comparisons. In an illiterate AlexNet trained on object recognition (Krizhevsky et al., 2017), the representations of these letters, also related to the Latin alphabet, help us contextualise the behavioural difference in a computational model of vision, lacking language information. The high dissimilarity between Braille on one side and Line Braille and Latin alphabet on the other can then explain the different transcription accuracies as a visual processing difference. In this view, in the absence of a strong linguistic mapping between scripts, the impact of visual features is more prominent and plays a more significant role in the response. Such role is likely dampened by top-down interactions during the learning of full words, where stronger linguistic associations between novel and consolidated script are made.

Similarly, the comparison between behavioural and computational learning curves expands this view to the later stages of learning. Both neural network models presented marked differences between Braille and Line Braille, diverging from the human results showing substantial similarities between scripts, which converge towards the end of the training. The distance between

these results can be attributed to the linguistic processing happening in human participants, who learnt to associate the words in the novel alphabet with existing lexical entries. Although we do simulate this process in computational models by adding Braille words to the same output category of Latin alphabet classes, human linguistic processing and its role in shaping visual areas are more nuanced than semantic associations (Dębska et al., 2024; Liu et al., 2025; Planton et al., 2022; S. Wang et al., 2022; X. Wang et al., 2018; Woolnough et al., 2021) and could not be modelled in our experiment. Overall, the letter dissimilarity and the accentuated learning delay found in computational models offer a view of a purely visual processing of Braille, which does not align with the increased similarity between Braille and Line Braille found in humans, largely to be attributed to linguistic processing.

The network architectures trained on Braille in combination with Latin alphabet can be considered an expert network and can be compared to the human participants from the fMRI experiment. Overall, as already discussed in chapter 4, we did not find an alignment between network and human representations. While Braille in expert human participants is processed akin to Latin alphabet, using similar coding principles as a result of expertise, the network clearly processes the scripts based on their visual features, with minimal effect of training. At the present moment, the analyses on the full visual hierarchy are not complete and the models behaviour in layers corresponding to V1, V4, IT, cannot be explored in detail. However, the increase in the mean dissimilarity across layers, in both networks, serves as an indication of a visual processing that differs from the multivariate decoding accuracies found in human brain regions. From the neural data, one would

expect an early distinction between familiar and unfamiliar orthographies, which we found as early as V1, while in the models' dissimilarities increase later in the hierarchy and with different patterns per script, more for Latin than for Braille across expertise. Moreover, the finer-scale multivariate organization at different stages further indicates that the network's representations are driven by visual features rather than expertise and linguistic content. This different processing can be reconducted to the absence of language processing seen in learning a script. The networks have no information about the linguistic properties of the words tested, apart from being mapped to an output class (which can parallel a lexical entry), meaning that the processing relies only on the statistical redundancies that can be extracted from the visual features. In this sense, a high dissimilarity between Latin alphabets and Braille is not surprising, given the respective presence and absence of line junctions in the Latin alphabet and in Braille. It appears then evident that the difference between networks' representations and human reading because the linguistic information influences the visual processing in multifaceted ways (Pattamadilok et al., 2019; Planton et al., 2022; S. Wang et al., 2022; X. Wang et al., 2018).

Once combined, these studies form a clear picture of the processing of visual Braille. The linguistic information of the letters and words is partially abstracted from the visual features and leads to neural responses that rely on the same coding principles as canonical scripts based on line junctions, albeit with minor differences between orthographies. This information can be used as a basis for a broader understanding of the reading network.

The behavioural and neural data presented in this dissertation can be used to predict the trajectory of Braille learning from behaviour to neural representations, a trajectory that can be abstracted to any visual way of conveying linguistic information, regardless of its basic visual features. Previous studies trained individuals on new orthographies, among which Braille, to look at the neural correlates of learning a script (Gaca et al., 2025; Martin et al., 2019; Moore et al., 2014; Siuda-Krzywicka et al., 2016) and found VWFA activation for the trained stimuli after different time windows. In particular, Gaca and colleagues trained novice participants on the learning of tactile Braille letters of progressive difficulty and observed VWFA activation after only seven days of training (Gaca et al., 2025). Even more interestingly, although the participants were trained on the tactile modality, the authors also observed activations for visual presentations of Braille letters in the same timeframe. This result is in line with what has been observed in participants after extensive training (Siuda-Krzywicka et al., 2016) and, given the short time window, allow us to infer what kind of functional adaptation in VWFA may be induced by the training on Braille and Line Braille. In this case, one could expect the same trend of results observed in the expert participants of our study (chapter 3), possibly modulated by the level expertise (i.e. the amount of learning). Participants should exhibit the same trend of similarity between their trained script (Braille or Line Braille) and their native Latin alphabet, with separate neuronal populations encoding the scripts, even in the case of Line Braille. Here, although both scripts are based on line junctions, a similar distinction from the Latin alphabet should be expected. Zhan and colleagues showed additional sub-patches of VWFA specific to Chinese in English-Chinese bilinguals (languages based on different orthographies) and not in English-

French bilinguals (languages based on the same orthography) (Zhan et al., 2023). Based on these results, additional or entirely separate sub-patches should emerge for different orthographies, including Line Braille or any script that differs from the Latin alphabet, as an effect of the visual modulation of linguistic organization that spans across the reading network.

The trained reading network seems indeed to tune to linguistic properties from its earliest visual stages. Linguistic information determines the response pattern of key areas in the visual stream when participants process words and word-like stimuli, even in the absence of a univariate preference for a script. These results support and expand a view of interactions between VWFA and the language network to other visual regions, in particular showing a preference for familiar scripts in V1 (Szwed et al., 2014) and evidence of processing of linguistic stimuli in LOC (Haupt et al., 2024). Even more remarkably, we observed representations similar to those determined by language even for Braille, when an individual was familiar with it, indicating a strong interaction between vision and language in reading (Price & Devlin, 2003, 2011) and in the general representation of knowledge (Bi, 2021; Liu et al., 2025). At the same time, it is worth addressing already that from the information in this dissertation we cannot yet exclude the perception of imaginary lines between Braille dots. We can however make inferences based on the presented results. In chapter 3, we explored the neural responses to visual Braille reading and found multivariate patterns of linguistic information in the visual stream (V1, LOC) in the absence of univariate preference for Braille, and we did not find significant cross-script decoding between line- and dot-based scripts. From the learning experiment, minor differences between

the scripts indicate different visual processing, which would not be present if imaginary lines were perceived in Braille, effectively creating the line junctions present in Line Braille. Adding to this, the representation of letters in computational models (chapter 4) shows marked dissimilarities between Braille and Line Braille. While such a result is not conclusive by itself, it expands the behavioural results by showing the purely visual feed-forward processing of these scripts, with and without lines. This aspect is further explored later among other limitations, to suggest a framework for future studies to address the questions experimentally.

Limitations and future perspectives

This dissertation elucidates the processes that influence reading and object categorization. However, some limitations in the studies remain and addressing them is important to provide further insights and guide additional future research.

The possibility that participants might perceive imaginary connecting lines between Braille dots can be seen as one of the main limitations of this corpus of research. Illusionary contours (Kanisza, 1976; Kennedy & Ware, 1978) may be cited as an argument in favour of the completion made from separate dots. In the absence of direct evidence in favour or against, an experiment can be suggested to clarify the matter. A starting experiment could be a study where participants are presented with two dots on a screen and (1) are also presented with a line connecting those dots, (2) are asked to imagine a line

connecting the dots, or (3) simply perceive the two dots, variating the distance and size of these dots. The comparison of neural activations in response to the different conditions could reveal whether conditions 1 and 2 differ or not from each other, and on what scale. Additionally, a sufficiently large number of conditions could help investigate the tipping point for such illusion to appear, for example a needed proximity of two dots in relation to their sizes. Successive, more direct studies should then expand this phenomenon in the naturalistic context of Braille and compare it to Line Braille as a control script, and possibly to additional custom scripts where the dots do not follow the grid arrangement, to control for factors that may facilitate the emergence of illusory lines. Such corpus of research should elucidate the processes of visual integration of Braille.

Another limitation of Braille concerns the temporal aspect of reading. While reading words in the Latin alphabet is a fast and holistic process (i.e. the whole word is perceived at the same time), in most cases Braille is instead a serial process of letter-by-letter reading, slower compared to canonical scripts. In this context, it cannot be fully excluded that different reading mechanisms lead to differences in the neural activations. In the design of the multivariate fMRI experiment, we opted for a long exposure time for both Latin alphabet and Braille, to standardize the stimulus length. While the strong results of Representational Similarity Analysis (RSA) highlight a highly similar organization and highly similar processing, excluding substantial differences in the processing of the two scripts, investigating more in depth the temporal evolution of visual Braille reading could better elucidate the steps of language access and their impact on reading. Woolnough and colleagues, using

intracranial recordings, unravelled two distinct gradients of activation that may discern the bottom-up processing of visual properties of the stimuli presented from the top-down influences of language (Woolnough et al., 2021). Similarly, the slower temporal dynamics of visual Braille reading could help separate these two processes and further explore this interplay between bottom-up and top-down streams of activation.

The spatial resolution of the fMRI experiment poses an additional limitation. Compared to recent studies, in our localization of VWFA (chapter 3) we did not identify multiple sub-regions but rather a broad activation. This result is not against the literature (Caffarra et al., 2021; Dębska et al., 2023; Lerma-Usabiaga et al., 2018; Pillet et al., 2024; Yeatman & White, 2021; Zhan et al., 2023), but is instead a limitation of the experimental design and the methodology. First, to probe for the response to Braille in the expert group we opted for two independent VWFA localizers, separately with Latin alphabet words and with Braille words. Due to the constraints of scanning time, we did not include additional contrasts that could have compared meaningful and non-meaningful stimuli in both alphabets to highlight sub-areas of VWFA. Now that activation related to visually presented Braille has been established, further studies can be designed to compare the responses of two sub-regions and to obtain a more comprehensive picture of VWFA. A careful design that functionally localizes VWFA sub-regions, or that defines the anatomical areas deputed to the perceptual and lexical VWFAs (Lerma-Usabiaga et al., 2018) from atlases, could probe for different responses within Braille. Building on the existing literature, one could expect a more posterior patch, tuned to the visual features, to show stronger differences between Latin and Braille scripts,

possibly even at the multivariate level. At the same time, a more anterior patch of VWFA has been found to have strong connections with the language network (Dębska et al., 2023; Lerma-Usabiaga et al., 2018; Yeatman & White, 2021). In this context, the responses for different alphabets in such sub-region could show even higher similarity and potentially converging and generalized representations at the multivariate level.

The current state of computational models of reading can well represent the level of visual features, presenting units selective for the characteristics or the position of letters (Agrawal & Dehaene, 2024), but does not align to higher level human representations of Braille and other scripts. While this lack of convergence provides useful information about the brain, further steps can be made to develop more comprehensive reading models, incorporating the processing of linguistic information. To build models that encompass the functioning of VWFA and of the reading network as a whole, it will be important to complement the visual processing layers with layers that elaborate the different linguistic aspects of words and pieces of text at a higher level, representing for example relationships between words and concepts. Vision-language models (VLM) can offer such solution, although this class of models has the different goal of objects or scene perception from two different modalities. Additionally, given evidence for processing of other sensory modalities in VWFA, the inclusion of speech processing network alongside the visual processing may provide stronger results. Moreover, a reliable model should account for the impairments found in humans. Testing the model's architecture by simulating deficits caused by learning disabilities or stroke-related (e.g. alexia) could lead to stronger models that could in turn

supplement neuroimaging and behavioural methods to investigate the typical and atypical functioning of VWFA and the reading network.

Lastly, this dissertation approached reading from its visual component, and focused on the results about language coming from minor or the absence of differences between Braille and the Latin alphabet. A further step forward can be done to separate visual features and linguistic properties. Previously, Vinckier and colleagues created a gradient of stimuli with different linguistic properties containing strings of false-font, of infrequent letters, of frequent letters but rare bigrams, of frequent bigrams but rare quadrigrams, of frequent quadrigrams, of real words (Vinckier et al., 2007). This stimulus set has been widely used or replicated in many studies (Cerpelloni et al., 2024; Woolnough et al., 2021; Zhan et al., 2023), due to its structure that allows to operationalize the statistical regularities present in the letters and words we read. However, it focuses on the visual component of the words but does not directly expand on the linguistic counterpart. Additional manipulations of the linguistic aspects of the stimuli, including for example graphemes that can be associated to different phonemes, could unravel differences in the visual processing of those graphemes based on language. Similarly, adjusting the meaning of the real words, and creating relations between them could tap into the organization of knowledge coming from a visual source. As seen in many studies, the interplay between language and vision in reading is evident, and a series of experiments that actively measures both components is needed to fully understand the impact of language in reading and generally in the visual system.

Conclusion

In this dissertation, I expand the literature on the Visual Word Form Area and on the reading network by employing the visual processing of Braille words in three complementary studies using methodologies ranging from behavioural experiments to deep neural networks. Through the main findings that learning to read a new script relies on the strength of the mapping between novel and consolidated words, that VWFA responses underlie similar structures related to the linguistic content, and that visual models do not possess the same characteristics, I demonstrate that humans integrate linguistic and visual information when reading. Overall, this dissertation shows that the recognition of words does not ignore their meaningful information, and that this content, linguistic in nature, drives the unique human ability that is reading.

References

Agrawal, A., & Dehaene, S. (2024). Cracking the neural code for word recognition in convolutional neural networks. *PLOS Computational Biology*, *20*(9), e1012430. https://doi.org/10.1371/journal.pcbi.1012430

Baker, C. I., Liu, J., Wald, L. L., Kwong, K. K., Benner, T., & Kanwisher, N. (2007). Visual word processing and experiential origins of functional selectivity in human extrastriate cortex. *Proceedings of the National Academy of Sciences*, *104*(21), 9087–9092. https://doi.org/10.1073/pnas.0703300104

Bi, Y. (2021). Dual coding of knowledge in the human brain. *Trends in Cognitive Sciences*, *25*(10), 883–895. https://doi.org/10.1016/j.tics.2021.07.006

Bola, Ł., Radziun, D., Siuda-Krzywicka, K., Sowa, J. E., Paplińska, M., Sumera, E., & Szwed, M. (2017). Universal Visual Features Might Be Necessary for Fluent Reading. A Longitudinal Study of Visual Reading in Braille and Cyrillic Alphabets. *Frontiers in Psychology*, *8*. https://doi.org/10.3389/fpsyg.2017.00514

Büchel, C., Price, C., & Friston, K. (1998). A multimodal language region in the ventral visual pathway. *Nature*, *394*(6690), 274–277. https://doi.org/10.1038/28389

Caffarra, S., Karipidis, I. I., Yablonski, M., & Yeatman, J. D. (2021). Anatomy and physiology of word-selective visual cortex: From visual features to lexical processing. *Brain Structure and Function*, *226*(9), 3051–3065. https://doi.org/10.1007/s00429-021-02384-8

Cerpelloni, F., Van Audenhaege, A., Matuszewski, J., Gau, R., Battal, C., Falagiarda, F., Op de Beeck, H., & Collignon, O. (2024). Expertise in reading visual Braille co-opts the reading network. *bioRxiv*, 2024.05.08.593104. https://doi.org/10.1101/2024.05.08.593104

Changizi, M. A., Zhang, Q., Ye, H., & Shimojo, S. (2006). *The Structures of Letters and Symbols throughout Human History Are Selected to Match Those Found in Objects in Natural Scenes.* 23.

Dębska, A., Wang, S., Jednoróg, K., & Pattamadilok, C. (2024). *Reading Acquisition Drives Linguistic Cross-Modal Convergence in The Left Ventral Occipitotemporal Cortex*. Neuroscience. https://doi.org/10.1101/2024.11.26.625405

Dębska, A., Wójcik, M., Chyl, K., Dzięgiel-Fivet, G., & Jednoróg, K. (2023). Beyond the Visual Word Form Area – a cognitive characterization of the left ventral occipitotemporal cortex. *Frontiers in Human Neuroscience*, *17*, 1199366. https://doi.org/10.3389/fnhum.2023.1199366

Dehaene, S., & Cohen, L. (2011). The unique role of the visual word form area in reading. *Trends in Cognitive Sciences*, *15*(6), 254–262. https://doi.org/10.1016/j.tics.2011.04.003

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

Gaca, M., Olszewska, A. M., Droździel, D., Kulesza, A., Paplińska, M., Kossowski, B., Jednoróg, K., Matuszewski, J., Herman, A. M., & Marchewka, A. (2025). How learning to read Braille in visual and tactile domains

reorganizes the sighted brain. *Frontiers in Neuroscience*, *18*, 1297344. https://doi.org/10.3389/fnins.2024.1297344

Haupt, M., Graumann, M., Teng, S., Kaltenbach, C., & Cichy, R. (2024). The transformation of sensory to perceptual braille letter representations in the visually deprived brain. *eLife*, *13*, RP98148. https://doi.org/10.7554/eLife.98148

Kanisza, G. (1976). Subjective Contours. *SCIENTIFIC AMERICAN*.

Kennedy, J. M., & Ware, C. (1978). Illusory Contours Can Arise in Dot Figures. *Perception*, *7*(2), 191–194. https://doi.org/10.1068/p070191

Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2017). ImageNet classification with deep convolutional neural networks. *Communications of the ACM*, *60*(6), 84–90. https://doi.org/10.1145/3065386

Kubilius, J., Schrimpf, M., Nayebi, A., Bear, D., Yamins, D. L. K., & DiCarlo, J. J. (2018). *CORnet: Modeling the Neural Mechanisms of Core Object Recognition*. Neuroscience. https://doi.org/10.1101/408385

Lerma-Usabiaga, G., Carreiras, M., & Paz-Alonso, P. M. (2018). Converging evidence for functional and structural segregation within the left ventral occipitotemporal cortex in reading. *Proceedings of the National Academy of Sciences*, *115*(42). https://doi.org/10.1073/pnas.1803003115

Liu, B., Wang, X., Wang, X., Li, Y., Han, Y., Lu, J., Zhang, H., Wang, X., & Bi, Y. (2025). Object knowledge representation in the human visual cortex requires a connection with the language system. *PLOS Biology*, *23*(5), e3003161. https://doi.org/10.1371/journal.pbio.3003161

Martin, L., Durisko, C., Moore, M. W., Coutanche, M. N., Chen, D., & Fiez, J. A. (2019). The VWFA Is the Home of Orthographic Learning When Houses

Are Used as Letters. *Eneuro*, *6*(1), ENEURO.0425-17.2019. https://doi.org/10.1523/ENEURO.0425-17.2019

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

Moore, M. W., Durisko, C., Perfetti, C. A., & Fiez, J. A. (2014). Learning to Read an Alphabet of Human Faces Produces Left-lateralized Training Effects in the Fusiform Gyrus. *Journal of Cognitive Neuroscience*, *26*(4), 896–913. https://doi.org/10.1162/jocn_a_00506

Pattamadilok, C., Planton, S., & Bonnard, M. (2019). Spoken language coding neurons in the Visual Word Form Area: Evidence from a TMS adaptation paradigm. *NeuroImage*, *186*, 278–285. https://doi.org/10.1016/j.neuroimage.2018.11.014

Pillet, I., Cerrahoğlu, B., Philips, R. V., Dumoulin, S., & De Beeck, H. O. (2024). The position of visual word forms in the anatomical and representational space of visual categories in occipitotemporal cortex. *Imaging Neuroscience*, *2*, 1–28. https://doi.org/10.1162/imag_a_00196

Planton, S., Chanoine, V., Sein, J., Anton, J.-L., Nazarian, B., Pallier, C., & Pattamadilok, C. (2019). Top-down activation of the visuo-orthographic system during spoken sentence processing. *NeuroImage*, *202*, 116135. https://doi.org/10.1016/j.neuroimage.2019.116135

Planton, S., Wang, S., Bolger, D., Bonnard, M., & Pattamadilok, C. (2022). Effective connectivity of the left-ventral occipito-temporal cortex during visual word processing: Direct causal evidence from TMS-EEG

co-registration. *Cortex*, *154*, 167–183. https://doi.org/10.1016/j.cortex.2022.06.004

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

Price, C. J., & Devlin, J. T. (2004). The pro and cons of labelling a left occipitotemporal region: "The visual word form area". *NeuroImage*, *22*(1), 477–479. https://doi.org/10.1016/j.neuroimage.2004.01.018

Price, C. J., & Devlin, J. T. (2011). The Interactive Account of ventral occipitotemporal contributions to reading. *Trends in Cognitive Sciences*, *15*(6), 246–253. https://doi.org/10.1016/j.tics.2011.04.001

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

Siuda-Krzywicka, K., Bola, Ł., Paplińska, M., Sumera, E., Jednoróg, K., Marchewka, A., Śliwińska, M. W., Amedi, A., & Szwed, M. (2016). Massive cortical reorganization in sighted Braille readers. *eLife*, *5*, e10762. https://doi.org/10.7554/eLife.10762

Szwed, M., Cohen, L., Qiao, E., & Dehaene, S. (2009). The role of invariant line junctions in object and visual word recognition. *Vision Research*, *49*(7), 718–725. https://doi.org/10.1016/j.visres.2009.01.003

Szwed, M., Dehaene, S., Kleinschmidt, A., Eger, E., Valabrègue, R., Amadon, A., & Cohen, L. (2011). Specialization for written words over objects in the visual cortex. *NeuroImage*, *56*(1), 330–344. https://doi.org/10.1016/j.neuroimage.2011.01.073

Szwed, M., Qiao, E., Jobert, A., Dehaene, S., & Cohen, L. (2014). Effects of Literacy in Early Visual and Occipitotemporal Areas of Chinese and French Readers. *Journal of Cognitive Neuroscience*, *26*(3), 459–475. https://doi.org/10.1162/jocn_a_00499

Van Audenhaege, A., Mattioni, S., Cerpelloni, F., Remi, G., Arnaud, S., & Collignon, O. (2024). *Phonological representations of auditory and visual speech in the occipito-temporal cortex and beyond*. https://doi.org/10.1101/2024.07.25.605084

Vinckier, F., Dehaene, S., Jobert, A., Dubus, J. P., Sigman, M., & Cohen, L. (2007). Hierarchical Coding of Letter Strings in the Ventral Stream: Dissecting the Inner Organization of the Visual Word-Form System. *Neuron*, *55*(1), 143–156. https://doi.org/10.1016/j.neuron.2007.05.031

Vogel, A. C., Petersen, S. E., & Schlaggar, B. L. (2012). The Left Occipitotemporal Cortex Does Not Show Preferential Activity for Words. *Cerebral Cortex*, *22*(12), 2715–2732. https://doi.org/10.1093/cercor/bhr295

Wang, S., Planton, S., Chanoine, V., Sein, J., Anton, J.-L., Nazarian, B., Dubarry, A.-S., Pallier, C., & Pattamadilok, C. (2022). Graph theoretical analysis reveals the functional role of the left ventral occipito-temporal cortex in speech processing. *Scientific Reports*, *12*(1), 20028. https://doi.org/10.1038/s41598-022-24056-1

Wang, X., Xu, Y., Wang, Y., Zeng, Y., Zhang, J., Ling, Z., & Bi, Y. (2018). Representational similarity analysis reveals task-dependent semantic influence of the visual word form area. *Scientific Reports*, *8*(1), 3047. https://doi.org/10.1038/s41598-018-21062-0

Woolnough, O., Donos, C., Rollo, P. S., Forseth, K. J., Lakretz, Y., Crone, N. E., Fischer-Baum, S., Dehaene, S., & Tandon, N. (2021). Spatiotemporal

dynamics of orthographic and lexical processing in the ventral visual pathway. *Nature Human Behaviour*, *5*(3), 389–398. https://doi.org/10.1038/s41562-020-00982-w

Xue, G., & Poldrack, R. A. (2007). The Neural Substrates of Visual Perceptual Learning of Words: Implications for the Visual Word Form Area Hypothesis. *Journal of Cognitive Neuroscience*, *19*(10), 1643–1655. https://doi.org/10.1162/jocn.2007.19.10.1643

Yeatman, J. D., & White, A. L. (2021). Reading: The Confluence of Vision and Language. *Annual Review of Vision Science*, *7*(1), 487–517. https://doi.org/10.1146/annurev-vision-093019-113509

Zhan, M., Pallier, C., Agrawal, A., Dehaene, S., & Cohen, L. (2023). Does the visual word form area split in bilingual readers? A millimeter-scale 7-T fMRI study. *Science Advances*.