

1. Resumen

El cerebro humano es fascinante, siendo capaz de procesar una gran cantidad de características pertenecientes a un rostro en cuestión de milisegundos, y sin aparentemente coste cognitivo. Este tipo de procesamiento parece ser inherente al ser humano y fundamental para múltiples áreas, entre ellas, la más evidente sería la socialización. Aunque el cerebro humano sigue siendo un gran desconocido, un hito histórico en el campo ha sido el descubrimiento de ciertas funciones del giro fusiforme. Concretamente, el reconocimiento de rostros se produce en la zona denominada FFA de dicho giro fusiforme, descrita como una de las regiones más especializadas para el reconocimiento facial en el sistema visual humano. Esta parte del cerebro también se ha relacionado con las tareas de reconocimiento de palabras, y de objetos, entre otros procesos. Sin embargo, la naturaleza del reconocimiento de palabras y caras parece ser muy compleja. Es, en este sentido, que debemos remarcar la posición del área visual de la forma de la palabra, popularmente denominada VWFA, también situada en el giro fusiforme izquierdo. La función específica del VWFA sería para gran parte de la literatura, el reconocimiento visual de cadenas de letras, dentro del procesamiento perteneciente al lector experto. No obstante, mientras el reconocimiento de caras humanas ha sido considerado innato, el reconocimiento de palabras debe ser aprendido. De hecho, invertimos gran parte de nuestra educación formal en este cometido.

De forma particular, algunos resultados, parecen apostar por diferencias individuales, tanto internas como externas. La literatura científica ha mostrado cómo ciertos factores relacionados con el lugar de residencia podrían ejercer un efecto en nuestro reconocimiento facial. Incluso el sexo, podría ser un motivo para las diferencias individuales. Obviamente, la casuística es marcadamente compleja, y se enfrenta a grandes retos, como el análisis del rol de la cultura, los medios de comunicación o incluso, la sociabilidad de la propia persona, entre otras variables. Otra variable de interés, dentro de los procesos internos, son los procesos atencionales y ejecutivos. Estos se encuentran relacionados con la forma en que se gestiona la información perceptiva. Por tanto, las diferencias en términos de procesos controlados o implícitos son de interés. Algunos

autores sugieren que el reconocimiento del rostro humano podría ser un proceso automático, así como, nuestro cerebro podría hacer uso de recursos atencionales independientes para los abordajes por características y de configuración. En este sentido, otros autores sugieren que el procesamiento holístico podría ser también automático. Si se comparan estos resultados con el procesamiento de los estímulos escritos, la literatura parece más explícita hacia una independencia atencional.

Por otro lado, el área del VWFA ha sido descrito como parte de los circuitos de lenguaje y atención. Sin embargo, como las formas de las palabras son más abstractas que las características específicas de otros estímulos, la invariabilidad de la posición podría apoyar la independencia atencional. Esta independencia también puede encontrarse cuando los estímulos se presentan fuera del foco de atención. Los aspectos atencionales y ejecutivos han sido estudiados a través de otros enfoques, como los componentes de latencias de respuesta. Recientemente, una investigación abordó el procesamiento de caras a través de un ajuste ex-Gaussiano, un análisis que justamente nos permite descomponer los tiempos de respuesta en componentes. Otra estrategia, podría incluir la diversificación en tipo de tareas, por ejemplo, tareas de reconocimiento simple, donde los efectos son más tardíos, o tareas de *priming*, donde los efectos se consideran más tempranos. La manipulación del subtipo de *priming* denominado enmascarado, es decir, el que incluye una máscara entre el estímulo *prime* y *target*, evidencia del procesamiento léxico automático y sin estrategia. De hecho, existe evidencia que sugiere que la naturaleza de los efectos encontrados en este paradigma podría ser pre-lexica.

El objetivo general de la presente tesis doctoral fue el estudio del procesamiento de caras a través de los componentes de respuesta. Para ello, se han realizado un total de cinco estudios compuestos por ocho experimentos. De forma particular, se establecieron un total de cuatro objetivos específicos. En primer lugar, y dado que no existe un corpus lingüístico para determinar la frecuencia de los rostros humanos, el primer objetivo instrumentales fue examinar si la frecuencia que proporcionan los motores de búsqueda puede considerarse como medida que refleje dicho efecto. Nuestra hipótesis se formula en este sentido de la siguiente forma: las frecuencias obtenidas a través de motores de búsqueda online predicen de forma similar los tiempos de reacción en tareas de decisión léxica que los corpus lingüísticos tradicionales. Tal y como marca la literatura previa con Blogs, Twitter y subtítulos de películas, este tipo de estímulos podría representar un entorno más ecológico de la vida cotidiana. Además, de demostrarse esto a través de la comparación con los corpus lingüísticos tradicionales en el reconocimiento de palabras

escritas, podrá desarrollarse el segundo objetivo. Tras confirmarse la hipótesis anterior, el segundo objetivo trató de explorar el efecto de frecuencia, tanto en el reconocimiento de palabras escritas, como en el de caras. Esto se llevó a cabo a través de la información proporcionada por los motores de búsqueda en línea, que, además, pueden proporcionar información relativa a las noticias, lo que, a su vez, podría considerarse un reflejo del contexto. Desde el enfoque anteriormente descrito, esto nos permitió evaluar tanto el efecto de frecuencia como de contexto, siempre que se asuma que las noticias sean un reflejo del contexto. Además, se examinó la posteridad de los efectos en el tiempo, a través de una réplica tras seis meses. Se hipotetizó, por tanto, que, la frecuencia de las búsquedas en línea y noticias presentan una mejor predicción de las latencias de reconocimiento de las palabras impresas frente a las de los rostros, y que, además, estas fluctúan en periodos breves de tiempo (6 meses).

El tercer objetivo versa sobre los aspectos atencionales, tratando de evaluar cuál se produce de forma más temprana, a través de un paradigma de *priming* enmascarado. Se hipotetizó que el efecto de *priming* enmascarado es más marcado para los estímulos escritos frente a los rostros. A través de un ajuste *ex-Gaussiano* se examinaron los valores de los parámetros en las diferentes tareas, así como, en participantes de diferentes países de proveniencia. El último objetivo, trató de evaluar la inestabilidad de los rostros humanos y los procesos de abstracción subyacentes, tal y como descrito en los modelos teóricos. Por ello, se propone un análisis de la ejecución de los participantes en, a diferencia de las tareas anteriores, el reconocimiento de rostros. El punto de partida para evaluar a los mejores participantes, tal y como marca la literatura, sería la capacidad de descartar nuevos rostros. Concretamente, se ha seleccionado el reconocimiento en diferentes niveles de posición. Además, dado que se ha encontrado evidencia que liga el tipo de escritura con los sesgos hacia derecha o izquierda, se hipotetizó un sesgo hacia la izquierda para los participantes lectors con dicho sesgo.

En lo que concierne a la metodología, se llevaron a cabo cinco estudios con un total de ocho experimentos, con diferentes muestras y submuestras. En el primer estudio, se examinó el efecto de frecuencia en línea a través de búsquedas en motores de búsqueda, en tareas de decisión léxica (*Go-No go*, y *Yes-No*). En el segundo estudio, se realizó una tarea de discriminación simple de caras y nombres de celebridades o personalidades reconocidas de forma internacional. En este sentido, se realizó un primer experimento en dos submuestras de estudiantes españoles y norteamericanos, así como un segundo experimento longitudinal bajo la misma tarea que examinaba los efectos *test-retest* tras 6

meses. En el tercer estudio, se trató de evaluar nuevamente los efectos del país de procedencia, así como controlar los efectos de sexo. Tanto en este experimento como en el número cuatro, se trató de estudiar la temporalización de los efectos, a través de un paradigma *priming* enmascarado. El cuarto estudio fue una réplica sistemática del estudio anterior, donde se examinaron los componentes ex-Gaussianos y concretamente el parámetro τ . Finalmente, en el quinto estudio se evaluó nuevamente los componentes ex-Gaussianos de latencias de respuesta a rostros humanos en diferentes posiciones, a través de una tarea de discriminación simple. Además, se ha desarrollado un *script* para interactuar con un API (*application programming interface*) que permita la obtención de frecuencias de búsqueda online (en el motor de búsqueda Google), y un segundo *script* para el cálculo de componentes ex-Gaussianos, ambos en lenguaje Python.

En el primer estudio, se empleó un conjunto de palabras de cinco letras, empleado en la literatura previa con adultos. La mitad de los estímulos eran de alta frecuencia y la segunda mitad era de baja frecuencia. Además, se seleccionaron un conjunto de no palabras o pseudopalabras, a los efectos de desarrollar una TDL. En este caso, los participantes fueron instruidos a identificar las palabras y descartar las pseudopalabras.

Para los estudios 2, 3, y 4, y a los efectos de emplear una tarea de categorización, se seleccionó una lista de celebridades o personas de referencia a nivel internacional. En una fase inicial, se seleccionaron tres muestras diferentes de estudiantes universitarios españoles, norteamericanos y brasileños, que indicaron determinadas celebridades. Tras una selección de las más frecuentes, se elaboró la batería de los estudios dos, tres y cuatro. Para las tareas finales, los participantes recibieron instrucciones de identificar a las celebridades o personalidades internacionales, a través de su nombre o rostro y descartando los estímulos desconocidos. En el estudio dos se emplearon tareas simples con medidas repetidas y en los estudios tres y cuatro se empleó una tarea de *priming* enmascarado.

En el quinto experimento se empleó la batería de reconocimiento de caras KDEF del *Karolinska institutet*. Esta proporciona estímulos con imágenes de diferentes expresiones y posiciones. En nuestro caso, cada estímulo fue repetido varias veces en cinco ángulos diferentes (perfil izquierdo, medio perfil izquierdo, central, medio perfil derecho, perfil derecho). Tras un bloque de presentación, los participantes fueron instruidos a identificar los estímulos previos y descartar los novedosos. Este fue un diseño de medidas repetidas.

Para el Desarrollo de los diferentes objetivos de esta tesis doctoral, se desarrollaron dos *scripts*. El primer *script* se empleó para recabar las frecuencias de las búsquedas en línea en el motor Google, el cual, ofrece una serie de API para que los desarrolladores programen sitios web que interactúan con las herramientas de Google. Para los diferentes lenguajes de programación, Google dispone de bibliotecas que interactúan con estas API. En la secuencia de comandos de Python desarrollada se ha usado la *API Client Library* para interactuar con el Motor de búsqueda personalizado de Google (CSE).

Así como el uso de Python puede ser bastante común para el caso anterior, la literatura parece ser más bien escasa en términos de cálculos con la función ex-Gaussiana en Python. Por ello, se desarrolló este segundo script, el cual se ha denominado ExGUtils (*ex-Gaussian Utilities*).

Los principales resultados obtenidos pueden enumerarse de la siguiente forma: i) las frecuencias de búsqueda, especialmente de noticias, proporcionadas por los motores de búsqueda como *Google*, son predictores de los tiempos de reacción en el reconocimiento de palabras escritas, pero no para el reconocimiento visual de rostros; ii) el reconocimiento de nombres escritos y rostros parece estable en el tiempo, pero las caras serían más susceptibles al aprendizaje; iii) a través de análisis de redes, se sugiere que el reconocimiento de caras y nombres escritos presenta preferencias por su misma naturaleza aunque ambos procesos se encuentren interconectados, y esta estructura se replica entre países; iv) Se han encontrado efectos tempranos en el reconocimiento más marcados para el reconocimiento de rostros que el de palabras escritas; v) Se ha hallado un mayor efecto de *priming* de repetición y un mejor ajuste ex -Gaussiano en el reconocimiento de palabras escritas frente al de caras; vi) es posible identificar perfiles de rendimiento en el reconocimiento visual de caras que difieren en su posición, con un sesgo hacia la izquierda.

Palabras clave: *reconocimiento de caras, reconocimiento de nombres, palabras escritas, componentes de procesamiento, FFA, VWFA*

2. Abstract

The human brain is fascinating, being able to process a large number of characteristics belonging to a face in a matter of milliseconds, and without apparent cognitive cost. This type of processing seems to be inherent to the human being and fundamental for multiple areas, among them, the most evident would be socialization. Although the human brain is still a great unknown, a historical milestone in the field has been the discovery of certain functions of the fusiform gyrus. Specifically, face recognition occurs in the so-called FFA area of the fusiform gyrus, described as one of the most specialized regions for facial recognition in the human visual system. This part of the brain has also been related to word and object recognition tasks, among other processes. However, the nature of word and face recognition appears to be very complex. It is, in this sense, that we must emphasize the position of the visual area of the word form, popularly called VWFA, also located in the left fusiform gyrus. The specific function of the VWFA would be for much of the literature, the visual recognition of chains of letters, within the processing belonging to the expert reader. However, while the recognition of human faces has been considered innate, word recognition must be learned. In fact, we invest much of our formal education in this endeavor.

In particular, some results seem to rely on individual differences, both internal and external. Scientific literature has shown how certain factors related to place of residence could have an effect on our facial recognition. Even sex could be a reason for individual differences. Obviously, casuistry is markedly complex, and faces great challenges, such as the analysis of the role of culture, the media or even the sociability of the person himself, among other variables. Another variable of interest, within the internal processes, are the attention and executive processes. These are related to the way perceptive information is managed. Therefore, the differences in terms of controlled or implicit processes are of interest. Some authors suggest that the recognition of the human face could be an automatic process, as well as, our brain could make use of independent attentional resources for approaches by characteristics and configuration. In this sense, other authors suggest that holistic processing could also be automatic. If these results are compared with the processing of written stimuli, the literature seems more explicit towards an attentional independence.

On the other hand, the area of the VWFA has been described as part of the language and attention circuits. However, since word forms are more abstract than the specific characteristics of other stimuli, position invariance may support attentional independence. This independence can also be found when stimuli are presented outside the focus of attention. The attentional and executive aspects have been studied through other approaches, such as the components of response latencies. Recently, an investigation approached the processing of faces through an ex-Gaussian adjustment, an analysis that precisely allows us to break down response times into components. Another strategy, could include diversification in type of tasks, for example, simple recognition tasks, where the effects are later, or priming tasks, where the effects are considered earlier. The manipulation of the priming subtype called masking, that is, the one that includes a more expensive one between the prime and target stimulus, evidence of automatic lexical processing and no strategy. In fact, there is evidence to suggest that the nature of the effects found in this paradigm may be pre-lexical.

The general objective of the present doctoral thesis was the study of the processing of faces through the response components. For this purpose, a total of five studies composed of eight experiments have been carried out. In particular, a total of four specific objectives were established. Firstly, and given that there is no linguistic corpus to determine the frequency of human faces, the first instrumental objective was to examine whether the frequency provided by search engines can be considered as a measure that reflects this effect. Our hypothesis is formulated as follows: frequencies obtained through online search engines similarly predict reaction times in lexical decision tasks than traditional linguistic corpora.

As previous literature with Blogs, Twitter and movie subtitles points out, this kind of stimulus could represent a more ecological environment of daily life. Furthermore, if this is demonstrated through comparison with traditional linguistic corpora in written word recognition, the second objective can be developed. After confirming the previous hypothesis, the second objective tried to explore the effect of frequency, both in the recognition of written words and faces. This was done through the information provided by online search engines, which, in addition, can provide information related to news, which, in turn, could be considered a reflection of the context. From the approach described above, this allowed us to evaluate both the effect of frequency and context, provided that it is assumed that the news is a reflection of the context. In addition, we examined the posterity of the effects over time, through a replay after six months. It was

hypothesized, therefore, that the frequency of online searches and news items are more predictive of recognition latencies of printed words versus those of faces, and that these also fluctuate over short periods of time (6 months).

The third objective deals with the attentional aspects, trying to evaluate which is produced earlier, through a masked priming paradigm. It was hypothesized that the effect of masked priming is more marked for written stimuli in front of faces. Through an ex-Gaussian adjustment the parameter values in the different tasks were examined, as well as, in participants from different countries of origin. The last objective was to evaluate aspects of instability of human faces that require processes of abstraction, as described in the theoretical models. Therefore, an analysis of the participants' performance in, unlike the previous tasks, face recognition is proposed. The starting point to evaluate the best participants, as marked by the literature, would be the ability to discard new faces. Specifically, recognition has been selected at different position levels. In addition, since evidence has been found linking the type of writing with right or left biases, a left bias was hypothesized for literacy participants with such bias.

As far as methodology is concerned, five studies were carried out with a total of eight experiments, with different samples and sub-samples. In the first study, the effect of online frequency was examined through search engines, in lexical decision tasks (Go-No go, and Yes-No). In the second study, a simple face and name discrimination task of internationally recognized celebrities or personalities was performed. In this sense, a first experiment was performed in two subsamples of Spanish and North American students, as well as a second longitudinal experiment under the same task that examined test-retest effects after 6 months. In the third study, we tried to evaluate again the effects of the country of origin, as well as to control for the effects of sex. Both in this experiment and in number four, we tried to study the temporization of the effects, through a masked priming paradigm. The fourth study was a systematic replica of the previous study, where the ex-Gaussian components were examined, specifically the parameter τ . Finally, in the fifth study the ex-Gaussian components of response latencies to human faces in different positions were again evaluated through a simple discrimination task. In addition, a script was developed to interact with an API (application programming interface) to obtain online search frequencies (in the Google search engine), and a second script to calculate ex-Gaussian components, both in Python language.

In the first study, a set of five-letter words, used in previous literature with adults, was used. Half of the stimuli were high frequency and the second half were low frequency.

In addition, a set of non-words or pseudo-words were selected for the purpose of developing a TDL. In this case, participants were instructed to identify the words and discard the pseudo-words.

For studies 2, 3, and 4, and for the purpose of employing a categorization task, a list of celebrities or international reference persons was selected. In an initial phase, three different samples of Spanish, North American and Brazilian university students were selected, which indicated certain celebrities. After a selection of the most frequent ones, the battery of studies two, three and four was elaborated. For the final tasks, participants were instructed to identify international celebrities or personalities, through their name or face and discarding unknown stimuli. Simple tasks with repeated measurements were used in study two and a masked priming task was used in studies three and four.

The fifth experiment used the KDEF face recognition battery from the Karolinska institutet. It provides stimuli with images of different expressions and positions. In our case, each stimulus was repeated several times in five different angles (left profile, half left profile, central profile, half right profile, right profile). After a presentation block, the participants were instructed to identify the previous stimuli and discard the new ones. This was a repeated measurement design.

For the development of the different objectives of this doctoral thesis, two scripts were developed. The first script was used to collect the frequencies of online searches in the Google engine, which offers a series of APIs for developers to program websites that interact with Google tools. For the different programming languages, Google has libraries that interact with these APIs. In the Python script developed, the Client Library API has been used to interact with the Google Custom Search Engine (CSE).

While the use of Python may be quite common for the above case, the literature seems to be rather scarce in terms of calculations with the ex-Gaussian function in Python. This is why this second script was developed, which has been named ExGUtils (ex-Gaussian Utilities).

The main results obtained can be listed as follows: i) search frequencies, especially news frequencies, provided by search engines such as Google, are predictors of reaction times in written word recognition, but not for visual face recognition; ii) written name and face recognition seems stable over time, but faces would be more susceptible to learning; iii) through network analysis, it is suggested that face and written name recognition have preferences by their very nature although both processes are interconnected, and this structure is replicated across countries; iv) Early effects on

recognition have been found to be more pronounced for face recognition than for written words; v) A greater effect of repetition priming and a better ex-Gaussian fit has been found in the recognition of written words versus faces; vii) It is possible to identify performance profiles in visual recognition of faces that differ in position, with a bias towards the left.

Keywords: face recognition, name recognition, written words, processing components, FFA, VWFA