

*Exploring the role of genetics and
environment on the neural substrates of word
and rule learning*

Author

Anna Basora Marimon

TESI DOCTORAL UPF/2015

Supervisors

Dra. Núria Sebastián Gallés

Dr. Miguel Burgaleta Díaz

Departament de Ciències Experimentals i de la Salut



**Universitat
Pompeu Fabra**
Barcelona

*A la meva família, i en especial
als meus pares*

ACKNOWLEDGEMENTS - AGRAÏMENTS

Primer que tot vull donar les gràcies als meus pares per l'educació i el sentit del treball que m'han transmès durant tots aquests anys i al meu germà per estar sempre al meu costat. També vull agrair als meus avis la transmissió de la seva saviesa i per ensenyar-me que, al final, un acaba sent el record que deix en els altres. Al meu marit li vull agrair la infinita paciència que ha tingut durant tot aquest temps i que a pesar dels entrebancs, mai hagi deixat de lluitar per construir un futur junts. Als meus sogres i cunyada els hi vull donar les gràcies per totes les converses i ànims que m'han donat en els moments més durs de la tesi.

Si una cosa tinc clara és que aquesta tesi no s'hagués pogut realitzar sense la presència de tres persones: els meus directors, Núria i Miguel i la meua companya i amiga Yu. A la Núria li he d'agrair tota la confiança que ha dipositat en mi, la complicitat de la nostre relació, els coneixements transmesos i la seva dedicació i entrega. A Miguel le tengo que agradecer sus enseñanzas, sus constantes mensajes de ánimo que tantísimo me han ayudado y sobretodo, sus drásticas fechas de entrega. I want to thank Yu, my partner in crime, for all her infinite help, her happiness and her exceptional personality that filled of smiles and joy my days in the lab.

També vull donar les gràcies a l'Albert Costa per fer que m'interessés pel món de la investigació i per donar-me l'oportunitat d'anar a Boston on vaig tenir una de les millor experiències de la

meva vida. A la Bego i a la Mireia Hernández, només dir-vos que sou les meves ídols i un exemple a seguir. A la Silvia, la Cristina, la Santa, la Florència i sobretot al Xavi per ser els meus àngels de la guarda. Sense vosaltres hagués caigut en l'infern dels cables, els virus, la burocràcia eterna i sobretot, els triggers que no apareixen per més petades que li donis a l'ordinador.

Una infinitat de gràcies a les meves amigues del lab, Alex, Alba, Ana, Judith, Júlia, Irene, Paola, Cristina, Loreto, Pallabi i Karla, sense oblidar els homes de l'harem, el Volker i el Jesús, amb qui he compartit penes, desgràcies però sobretot bons i divertits moments. Sense elles i ells segur que no hagués pogut acabar aquesta tesi. I no em puc oblidar de les floretes del meu jardí: les de Reus, Aida, Conxi, Maria i Mireia i les de Barcelona, Anna, Helena, Marta, Rita, Sara, Clara i Ruben les i el qual durant més de 10 anys han estat un dels meus pilars existencials. Gràcies per ser part del que sóc avui.

Finalment, agrair a la meva àvia el haver-me dit un dia que...
“querer es poder”.

ABSTRACT

Individual differences have been observed in word and grammar learning. However, the genetic and environmental origin of these individual differences is unknown. Two separate studies addressed this issue. In the first study, monozygotic and dizygotic twins were exposed to a word and rule learning task while event-related potentials were recorded. Brain signal similarity between twins was measured, and indicated a weak influence of genetics in both processes. In the second study, white matter integrity estimates were obtained for language-relevant connections in a sample of monozygotic twins that also underwent a statistical word learning task. Twin differences in both white matter integrity and word learning performance were computed and correlated. Significant results in the left arcuate and inferior fronto-occipital fasciculus revealed that the environmental pressures affecting the integrity of these two structures partially explain individual differences in word learning. Taken together, both studies suggest that life experiences play a pivotal role in shaping the neural substrates of word and rule learning.

L'origen ambiental i genètic de les diferències individuals en l'aprenentatge de paraules i regles gramaticals és desconegut. Per aquest motiu es van realitzar dos estudis. En el primer, dos grups de bessons, monozigòtics i dizigòtics, varen desenvolupar una tasca d'aprenentatge de paraules i regles. Es va mesurar la similitud entre bessons en l'activitat cerebral i els resultats indicaren que els factors genètics eren poc rellevants. En el segon estudi es va obtenir la integritat de la matèria blanca d'un grup de bessons monozigòtics, els quals també varen realitzar una tasca d'aprenentatge de paraules. Es va mesurar la diferència entre bessons tant en la integritat de la matèria blanca com en el nivell d'aprenentatge i es van correlacionar ambdues mesures. Els resultats mostraren que els factors ambientals que afecten als fascicles arcuat i front-occipital esquerres explicarien, en part, les diferències individuals en l'aprenentatge de paraules. En conjunt, ambdós estudis suggereixen que les experiències juguen un paper central en el modelatge dels substrats neuronals associats a l'aprenentatge de paraules i regles.

PREFACE

Knowing more than one language has become an essential tool to make life both easier and more successful. There is no doubt that the new era of technology, where both communication and transportation are at hand for almost everyone living in a developed country, has promoted the necessity and/or the desire to know other languages. Moreover, due to the globalization we are experiencing, companies' staff selection usually requires employees to be skilled in more than one language. Hence, being successful in your job may also be dependent upon the acquisition of a new language.

The need of young generations to become bilingual or multilingual has put second language (L2) learning at the top of the research agenda compared to some decades ago (De Groot & Kroll, 2014). Indeed, questions such as *what are the basic mechanisms underlying the learning of a new language?* and *what factors contribute to attain a proficient level?* are of utmost importance for learners, educators and researchers.

One frequent approach to these questions has been to compare bilinguals' performance while varying some relevant factors, such as age of acquisition or amount of exposure. However, fully matching participants in these factors is usually difficult. Research on how infants solve the problem of language learning has pointed to the existence of very basic mechanisms which help them discover the words and the grammatical rules of the language of their

environment. In the present dissertation, we will approach the problem of individual differences in language learning by analyzing the individual differences in adult populations in these basic mechanisms which are fundamental for language acquisition.

When a language is heard for the first time, it is perceived as a sequence of concatenated speech sounds without any acoustic cue signaling the beginning or the end of each word. This is an important handicap when compared to written language, where blank spaces separate words most of the time. Therefore, one of the first steps that learners must face in order to master a new language, in particular in natural learning contexts, is the segmentation of the speech stream into the words it contains. In order to do that, it has been suggested that learners make use of a domain-general ability called statistical learning, which allows them to compute the probability of finding two syllables concatenated. As syllables appearing next to each other are more likely to be part of the same word, whereas syllables that rarely co-occur tend to belong to two different words, the computation of probabilities between syllables is a useful cue that learners exploit to segment speech into words (Fiser & Aslin, 2001, 2002; Saffran, Aslin, & Newport, 1996). Once speech segmentation is accomplished, these new elements can be kept in working memory (Lopez-Barroso et al., 2011) to be associated with a specific meaning (Graf Estes, Evans, Alibali, & Saffran, 2007) and saved in long-term storage (Reber, 2013). A second mechanism, namely computing “non-adjacent” regularities, would underlie the learning of grammatical structures (Peña,

Bonatti, Nespor, & Mehler, 2002) – although some controversy exists concerning the need of two separate mechanisms (e.g., Bonatti, Peña, Nespor, & Mehler, 2006; Perruchet, Peereman, & Tyler, 2006; Perruchet, Tyler, Galland, & Peereman, 2004). The computation of non-adjacent dependencies would allow listeners to learn, for instance, subject-verb agreement in languages like English.

As for the question of the factors contributing to better language learning, it has been shown that age of acquisition, amount of exposure, and motivation are crucial when defining how skilled a person is in an L2 (Bialystok & Hakuta, 1999; DeKeyser, 2000; Flege, Bohn, & Jang, 1997; Johnson & Newport, 1989; Weber-Fox & Neville, 1996). Moreover, innate predispositions (“aptitudes” or “talents”) also contribute to the successful learning of a new language. Indeed, different studies have reported important differences between individuals exposed to an L2 with equivalent age and amount of exposure, and presumably motivation (Sebastián-Gallés & Baus, 2005; Tsukada et al., 2005). Hence, these results suggest that genetic endowment also may play an important role in L2 learning ability.

In contrast with the amount of studies focused on the environmental causes of individual differences in L2 learning (Bialystok & Hakuta, 1999; DeKeyser, 2000; Flege et al., 1997; Johnson & Newport, 1989; Weber-Fox & Neville, 1996), little research is found in regard with its genetic sources (Dale, Harlaar, Haworth, &

Plomin, 2010). Indeed, most of the language-related genetic studies are centered on language development and language disorders (e.g., (Haworth, Davis, & Plomin, 2013; Hayiou-Thomas, Dale, & Plomin, 2012; Kang & Drayna, 2011; Newbury & Monaco, 2010). A significant amount of these studies have used standardized paper and pencil tests which, regardless of their advantages (e.g., easy data collection in large samples and at low costs), cannot provide fine-grained information about the basic processes that acquiring a new language entails, let aside its neural substrates (e.g., Byrne et al., 2009, 2013; Dionne, Dale, Boivin, & Plomin, 2003; Yulia Kovas et al., 2005).

The main goal of the present dissertation is to contribute to a better understanding of the origin of individual differences in language learning. More specifically, we aim at investigating the role of genetics *and* environment on two basic mechanisms underlying word segmentation and rule learning. We will use methods allowing the analysis of both function and structure of the brain.

This dissertation is organized into three sections. The first section includes the review of the literature along two axes. First, it reviews studies investigating word segmentation and rule extraction through the computation of distributional regularities. Second, it presents studies investigating the role of genetic and environmental factors on language learning. The section closes describing the specific goals, hypotheses and methodologies employed in this dissertation.

The second section includes two manuscripts that have been submitted to peer-reviewed journals. The first one is a twin study in which we investigated to what extent genetic and environmental factors contribute to the variance observed in the electrophysiological brain activity elicited during word and rule learning. The second manuscript is also a twin study, but in this case, our goal was to unveil how environmental pressures may account for the recently observed association between word learning and specific white matter (WM) structures.

The last section comprises the general discussion and the conclusions. It summarizes the main findings from the two studies, and it discusses the implications to the language learning literature, including how the methods we used can contribute to future twin studies. Finally, I will present possible future lines of research.

CONTENTS

	Pag.
<i>Acknowledgements</i>	<i>iv</i>
<i>Abstract</i>	<i>viii</i>
<i>Preface</i>	<i>x</i>
<i>Figures and tables</i>	<i>xviii</i>

1. THEORETICAL BACKGROUND..... 1

1.1. Language learning: the acquisition of words and rules..... 1

1.1.1. Word acquisition: the statistical learning mechanism..... 2

1.1.1.1. Word and statistical learning in language development..... 4

1.1.1.2. Word and statistical learning in adult populations..... 6

1.1.2. Rule acquisition and distributional information..... 9

1.1.3. Statistical and rule learning in non-language domains and in non-human animals..... 11

1.1.4. Neural substrates of word and rule learning..... 13

1.1.4.1. ERP components in word and rule learning..... 14

1.1.4.2. Neuroimaging studies of word and rule learning..... 18

1.2. Factors contributing to individual differences in second language learning..... 24

1.2.1. Environmental factors influencing language learning.... 25

1.2.2. Twin and genetic studies related to language learning... 27

1.2.2.1. The relevance of twin studies..... 28

1.2.2.2. Twin studies in language disorders..... 34

1.2.2.3. Language twin studies in healthy population.....	36
1.3. <i>The present work: goals, procedures and hypothesis</i>	40
2. EXPERIMENTAL SECTION I.....	45
<i>Manuscript 1. Low heritability of individual differences in word and rule learning: A twin study of Event-related Potentials</i>	45
3. EXPERIMENTAL SECTION II.....	75
<i>Manuscript 2. Environmental contributions to the white matter structures associated with word learning</i>	75
4. GENERAL DISCUSSION.....	105
4.1. <i>Summary of results and discussion</i>	106
4.1.1. Heritability in both word and rule learning.....	106
4.1.2. Environmental contribution to brain structures associated with statistical language learning.....	111
4.2. <i>Implications on the language learning literature</i>	115
4.2.1. Individual differences in second language learning.....	115
4.2.2. Word and rule extraction: one or two mechanisms?.....	117
4.3. <i>New methodological approaches</i>	119
4.4. <i>Future lines of research</i>	121
5. SUMMARY AND CONCLUSIONS.....	125
6. REFERENCES.....	127
ANNEX.....	153
<i>Exploring the genetic and environmental influence in both native and non-native phoneme perception</i>	153

FIGURES AND TABLES

	Pag.
Figure 1. Results reported in Peña et al. (2002).....	11
Figure 2. Results reported in Sanders et al. (2002).....	15
Figure 3. Results reported in De Diego-Balaguer et al. (2007).....	17
Figure 4. Results reported in McNealy et al. (2006).....	20
Figure 5. Results reported in López-Barroso et al. (2013).....	21
Figure 6. Examples of results from ACE analysis.....	31
Figure 7. Results reported in de Geus et al. (2007).....	34
Figure 8. Summary of the procedure and results for the first study.....	108
Figure 9. Summary of the procedure and results for the second study..	112

Manuscript 1

Figure 1. Example of two wavelet maps from a pair of twins.....	60
Figure 2. Percentage of accuracy for the Recognition phase.....	63
Figure 3. ERP grand average from the Learning phase.....	64
Figure 4. Average intra-pair maximum cross-correlation.....	66
Table 1. Stimuli used for each language.....	56

Manuscript 2

Figure 1. Percentage of accuracy for each language.....	83
Figure 2. Correlations between white matter and word learning.....	84
Figure 3. Representation of the white matter tracts.....	93
Table 1. Stimuli used in the two word learning tests.....	91

Annex

Figure A1. ERP grand average for native and non-native condition.....158

Figure A2. Average intra-pair maximum cross-correlation.....159

Table A1. MMN *t* test for native and non-native condition.....157

1. THEORETICAL BACKGROUND

1.1. Language learning: the acquisition of words and rules

A significant amount of research has focused on discovering how infants can acquire a language in an astonishing short amount of time and without any guidance. Research in the past 20 years has shed light on infants' learning mechanisms that allow them to organize the enormous amount of information coming from their environment. More specifically, they showed that at the age of 8 months infants can extract statistical information from the speech signal to segment it into its lexical units (Saffran et al., 1996). This ability would be essential for word consolidation. Other studies have indicated that infants also make use of algebraic-like computations to learn the rules in which grammar is based (Marcus, Vijayan, Bandi Rao, & Vishton, 1999). Therefore, it seems that infants are bequeathed with powerful learning mechanisms based on the decoding of the speech signal by extracting its underlying statistical patterns.

From these first findings on infants' use of statistical computations for word segmentation and rule learning, researchers have come up with other relevant questions for our research goals, in particular, whether adults use statistical computations as infants, whether these computations are specific to language and human species and what

neural underpinnings are associated to the learning of both words and rules. In the following sections I will review some relevant studies addressing these issues. The review is not intended to be comprehensive of all the studies performed so far, but to provide some background information.

1.1.1. Word acquisition: the statistical learning mechanism

Generally speaking, statistical learning refers to the capture of regularities in the environment, whether they be visual, auditory or even tactile (Conway & Christiansen, 2005; Kirkham, Slemmer, & Johnson, 2002; Saffran, Johnson, Aslin, & Newport, 1999). This type of learning can go from a simple frequency-counting task (Maye, Weiss, & Aslin, 2008; Maye, Werker, & Gerken, 2002), throughout associative sequences as we find in operant conditioning (Wright & Rivera, 1997) to a more complex and sophisticated probabilities, such as the ones from Bayesian statistics (Griffiths & Kalish, 2007). Likewise, the diversity of stimuli over which these patterns are established is also considerably extensive. In the auditory domain, the types of stimuli studied go from tones to syntactic structures (Saffran et al., 1999; Thompson & Newport, 2007) while in the visual domain, they go from simple geometric shapes like squares or triangles to faces or high-order spatial structures (Fiser & Aslin, 2001, 2002).

In the language field, “statistical learning” has a more specific sense, that is, it defines the ability to track sequential patterns contained in speech. Importantly, as opposed to a written text, where white spaces are found between words, speech does not have any auditory cue limiting word boundaries. Thus, another type of cue (i.e. sequential patterns) must be followed to know where the beginning and the end of each word is set. These patterns are called transitional probabilities (TPs), and they are used to segment speech into its lexical elements (Goodsitt, Morgan, & Kuhl, 1993; Miller & Selfridge, 1950; Saffran et al., 1996). TPs are defined as the probability that the element X will be followed by the element Y in the sequence XY (i.e., $Prob\ Y/X = Prob\ XY/Prob\ X$). Both X and Y can be phonetic segments, syllables or even words. For instance, in the sequence *cleverbrother*, *clever* and *brother* are words but *verbro* is not. The origin of this knowledge relies on the fact that the probability to find the syllable *ver* after the syllable *cle* and the syllables *ther* after the syllable *bro* is higher than the probability to find the syllable *bro* after the syllable *ver*. Hence, by capturing TPs, babies and L2 learners can accomplish speech segmentation.

However, it is important to mention that in natural speech, it is considerably difficult to control for all the factors responsible for language learning. For instance, when words like *the* and *a* are the first items of a sentence or clause, an intonational pattern as well as a reduction in stress are delimiting a pause in the speech sequence. Thus, to test how much both infants and L2 learners can rely on TPs (or other types of information present in the speech signal),

researchers have used artificial languages for a better control of the stimuli (Gómez & Gerken, 2000). One advantage of artificial languages is that learners do not have any previous knowledge about them, thus everyone's learning starts from scratch.

In the following three sections, I will present some relevant studies on this topic and how research has developed since its first findings.

1.1.1.1. Word and statistical learning in language development

The first infant study testing the ability to track statistical regularities in speech was carried out by Saffran and colleagues in 1996. They exposed two groups of 8-month-old infants during two minutes to a continuous sequence of syllables containing four trisyllabic words (e.g. *pabikutibudogolatudaropi...*) without the presence of cues defining word boundaries except for the statistical properties of the syllable sequences. After exposure to the language, word learning was tested using two types of contrasts: a word (e.g., *pabiku*) versus a part-word (e.g., *budogo*) (two syllables belonged to a word and the third belonging to a different word) and a word versus a non-word (e.g., *dobuku*) (three syllables from the speech stream but not in the order in which they appeared). Results showed that in both cases, infants preferred words compared to part- or non-words. Therefore, this study showed that infants were sensitive to statistical information (i.e., longer looking time towards part- and

non- word trials), thus demonstrating that they can discriminate between syllable sequences with different transitional probabilities.

Although this study was interpreted as showing a mechanism used to discover word boundaries, in fact, what it actually demonstrated was that infants could discriminate between sequences of sounds with different internal coherence. For that reason, Saffran (2001) and Graf Estes et al. (2007) addressed this issue with different approaches. In Saffran (2001) two groups of infants were confronted with either English or nonsense sentences containing words or part-words (e.g., “I like my *tubido*” or “zy fike ny *tubido*”) after completing a segmentation task (Aslin, Saffran, & Newport, 1998; Saffran et al., 1996). The results showed that infants exposed to English frames listened significantly longer to words than part-words in contrast with the nonsense sentence condition in which infants treated both words and part-word equally. Graf Estes et al. (2007) asked whether 17-month-old infants could associate meanings to the segmented words. In this case, infants, after being exposed to an artificial language stream, performed a label-object pairings test. For half of the subjects, the labels were words belonging to the artificial language heard at the beginning of the experimental session and for the other half, the labels were either non-words or part-words. Results showed that only infants in the word condition showed label-object learning. Summarizing, these studies showed that infants consider words segmented through statistical computations as possible lexical candidates in their native language.

There is no doubt that artificial languages have been significantly useful to discover many aspects of language acquisition. However, for the same reason they allow for the isolation of variables of interest, one potential criticism could be made regarding their lack of ecological validity. Pelucchi, Hay, & Saffran (2009) performed a statistical language learning experiment where they presented sentences in Italian to infants learning English who have never heard Italian before. The results showed that they could discern between familiar versus new Italian words as well as high- versus low-TPs words. Thus, this study demonstrates that statistical learning is not only restricted to simple stimuli, but can be applied to more complex and naturalistic stimuli.

To sum up, the studies presented in this section show that infants make use of the distributional information contained in the speech signal to segment it into units considered by infants as word candidates. Moreover, they showed that artificial languages are an extremely useful method to study language learning in infant populations.

1.1.1.2. Word and statistical learning in adult populations

Another important field in language research has focused on the role that statistical learning may play on word segmentation and

learning when adults are exposed to a new language. Saffran, Newport, & Aslin (1996) were the first to report the use of statistical computations in word segmentation by adults and follow up studies showed that adults, as infants, may only need two minutes to accomplish a word segmentation task (Peña et al., 2002).

Nevertheless, the comparison between infant and adult results is not straightforward. For instance, adults have previous knowledge about phonetics, phonotactics, lexicon, grammar and syntax of their first language (L1), which could either interfere or help the learning of the new language (Cunillera et al., 2009; Cunillera, Toro, Sebastián-Gallés, & Rodríguez-Fornells, 2006; Finn & Hudson Kam, 2008; Toro, Pons, Bion, & Sebastián-Gallés, 2011; Wang & Saffran, 2014). On one hand, it has been observed that previous experience with L1 phonological constraints intrudes statistical word segmentation in both Catalan (Toro et al., 2011) and English speakers (Finn & Hudson Kam, 2008) when L1 phonotactic rule violations were present in the speech stream. On the other hand, previous vocabulary knowledge (Cunillera et al., 2009) or even bilingualism (Wang & Saffran, 2014) may help word segmentation.

Another difference between the way infants and adults may face statistical learning refers to the maturation of the neural substrates (and thus processing efficiency). While one may expect an improvement of the statistical learning ability across ages as observed in other language abilities (Nelson, de Haan, & Thomas, 2012), the results from several statistical learning studies do not

show it (Kirkham et al., 2002; Saffran et al., 1999; Saffran, Newport, Aslin, Tunick, & Barrueco, 1997). They reveal an equitable performance between adults and infants (Saffran et al., 1999), adults and children (Saffran et al., 1997) and infants with different months-of-age (Kirkham et al., 2002). However, a more recent study (Arciuli & Simpson, 2011) found that, by manipulating the age of participants (from 5 to 12 years old) and the speed of stimuli presentation (from 800 ms to 200 ms), developmental differences can emerge.

Finally, another characteristic differentiating infant and adult learning is related to the differences between implicit and explicit learning. Even if studied in a laboratory, infants do not know that they are being tested, whereas children and adults can or cannot be aware of it, depending on the experimental procedure. However, at present, there is conflicting evidence regarding this issue (see Stevens, Arciuli, & Anderson, 2014 and Arciuli, Torkildsen, Stevens, & Simpson, 2014 for more specific information).

Wrapping up, studies with adults have shown that, as infants, they possess a statistical learning mechanism that helps them segment and learn words. However, since adults have previous knowledge about the language structure, adults' word segmentation could be performed differently from infants'. It has yet to be fully determined how many differences in brain maturation affect such basic processing.

1.1.2. Rule acquisition and distributional information

There is no doubt that the acquisition of the lexicon is not sufficient to master a language. The extraction of grammatical structures is necessary to know how words have to be combined and modified to construct meaningful sentences. That is, soon after statistical learning was considered a potential mechanism for word segmentation, Marcus et al. (1999) revealed that 7-month infants were able to learn algebraic-like structures through abstract computations. Importantly, they proposed the existence of an alternative mechanism necessary for the acquisition of grammar. In the subsequent years, a plethora of studies have investigated the type of dependencies that can be learned and the computations required to perform them (Sandoval & Gómez, 2013). Here we will only focus on one of the first studies of such nature, as it has been crucial for follow-up studies including our own. Peña et al. (2002) asked whether transitional probabilities between non-adjacent elements, such as the ones found in subject-verb agreement, could be extracted as fast as adjacent ones. They exposed participants to a 10-minute speech stream formed by words following an A_iXC_i rule in which the element A always predicted C (TP of 1) while the element X could vary among three different items (TP of 0.33). Thus, the speech stream was formed by 9 words belonging to three different family rules (e.g., *puliki*, *puRaki* and *pufoki*). After exposure, participants faced a two-alternative forced choice test where part-words and words were presented. The same procedure

was followed in a second experiment, but in this case using part-words and rule-words (i.e., words in which the X syllable appeared in the stream, but never between A_i and C_i). In the first experiment, participants performed above chance, whereas they failed in the second test. These findings suggested that participants did not use the same mechanism for the extraction of generalizations as for the computation of adjacent dependencies. To confirm this hypothesis, they proposed changing the nature of the signal by adding short pauses (25 ms) between words, leaving participants without the workload of word segmentation. The results obtained after such manipulation were positive (*see figure 1*). Just as it had been shown for word segmentation, participants only needed two minutes to trigger the grammatical-like computations to make generalizations. Thus, in agreement with Marcus et al. (1999), they also suggest the existence of two different mechanisms, one to discover the words and one to learn the structural regularities governing them.

Other studies have further investigated whether the use of additional cues similar to the ones used in Peña et al. (2002) (i.e., short pauses between words) would help the extraction of statistical structures. It has been shown that the presence of prosodic cues (Langus, Marchetto, Bion, & Nespors, 2012), the repetition of one of the edged elements of the structure to be segmented (Endress, Scholl, & Mehler, 2005) or an increase in the variability of the middle element (Gómez, 2002) are crucial for the extraction of rules. These studies support and complement Peña et al. (2002)'s findings.

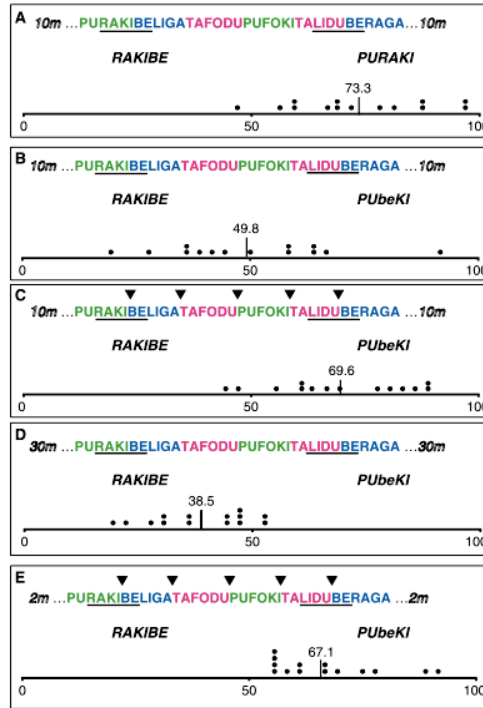


Figure 1. Results from the five experiments reported in Peña et al. (2002). (A) The segmentation test of an AXC 10-min stream, (B) of a rule-finding test, (C) of the same test over a stream with pauses, (D) without pauses but 30 min of exposure, and (E) 2 min of exposure with pauses. *Reproduced from Peña et al. (2002).*

1.1.3. Statistical and rule learning in non-language domains and in non-human animals

An important characteristic of statistical and rule learning is their domain generality. In particular, researchers wondered whether these mechanisms could be used with other types of auditory stimuli such as tones (Creel, Newport, & Aslin, 2004; Daikoku, Yatomi, & Yumoto, 2014; Endress, 2010; Saffran et al., 1999). Saffran et al. (1999) designed an experiment following the same procedure as in Saffran et al. (1996), but employing tones with different

frequencies. They showed that both adults and infants were capable of identifying word-tones (sequence of tones with high TP) from a tone stream. Subsequent studies investigated whether statistical learning would work in other modalities such as vision (Kirkham et al., 2002; Turk-Browne, Jungé, & Scholl, 2005), and they revealed that infants as well as adults were able to extract visual patterns.

Several studies have also investigated the domain-specificity of the learning of abstract rules. Marcus, Fernandes, & Johnson (2007) investigated whether 7.5 month-old infants could extract algebraic-like rules by using stimuli such as tones, instrument timbres and animal sounds. The results showed that infants extracted the rules only when they were previously exposed to a speech stream composed of the same rules. Similarly, Creel et al. (2004) and Endress (2010) showed that in order to capture the non-adjacent regularities in a sequence of tones, some acoustic cue (i.e., increase of pitch or timber among non-adjacent elements) or a specific context (i.e., non-adjacent elements are implemented in tonal melodies) must be present. Studies on visual rule learning reported that at the age of 7 months infants are aware of algebraic-like patterns (Johnson et al., 2009; Saffran, Pollak, Seibel, & Shkolnik, 2007) whereas adults can extract more complex visual regularities (Fiser & Aslin, 2001, 2002), being either temporal or spatial.

Another characteristic of both statistical learning and rule learning is its non-species specificity. That is, other species such as cotton-top monkeys (Hauser, Newport, & Aslin, 2001) or rats (Toro,

Trobalon, & Sebastián-Gallés, 2005) can extract statistical patterns from speech stimuli. However, the latter ones seem to have some limitations when compared to humans. Other studies investigating the extraction of rules in non-human primates (Newport, Hauser, Spaepen, & Aslin, 2004) and rodents (Murphy, Mondragón, & Murphy, 2008) have provided evidence showing that both similarities and differences exist across species.

Summarizing, these studies demonstrate that both statistical and rule learning are mechanisms that can be found in several domains and distinct modalities. Moreover, they are not only present in humans, but also in other mammals like monkeys and rats.

1.1.4. Neural substrates of word and rule learning

Once researchers proposed the existence of statistical and rule-learning mechanisms by using behavioral tests, they started asking how these types of computations were processed in the brain. To this end they combined the use of electroencephalography (EEG), as well as structural and functional magnetic resonance imaging (fMRI), including diffusion tensor imaging (DTI). In this section we first review studies using event related potentials (ERPs) and then studies applying neuroimaging techniques.

1.1.4.1. ERP components in word and rule learning

It was Sanders, Newport, & Neville (2002) who performed the first statistical learning study using ERPs. Participants in this study were evaluated on a word recognition test before and after training with an artificial language stream, similar to the one employed by Saffran et al. (1996). The sample was divided into high and low learners, based on the learning improvement, and ERPs from these two groups were compared before and after exposure. The results from this comparison showed an effect on the N1 (a negative amplitude enhancement at 100 ms after word presentation) only in the high learners group, as well as an increase of the N4 (a negativity at 400 ms after word onset) in high and low learners. The authors suggested that the increase in the N1 could be considered as an online measure signaling speech segmentation, whereas the increase in the N4 would indicate a lexical search process after word segmentation (*see figure 2*).

Subsequent studies also found a significant increase of both the N1 and the N4 when the neural activity elicited during a statistical learning condition was compared to a condition in which syllables were ordered randomly (Cunillera et al., 2009, 2006), as well as when high versus low learners were contrasted (Abla, Katahira, & Okanoya, 2008). Thus, these first ERP studies provided converging evidence of specific electrophysiological waves reflecting successful word segmentation in artificial languages.

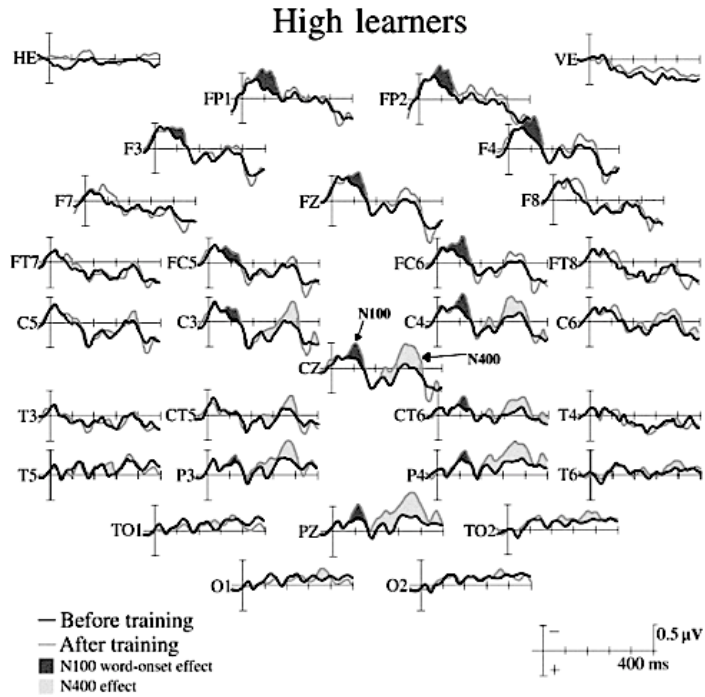


Figure 2. ERP results reported in Sanders et al. (2002). ERPs averaged to word onsets before and after training for the subjects showing the largest behavioral learning effects (high learners). After training, word onsets elicited a larger N100 at midline and medial electrode sites. Words also elicited a larger N400 after training. *Reproduced from Sanders et al., 2002.*

As said before, some authors posit different mechanisms for word and rule learning. De Diego Balaguer, Toro, Rodriguez-Fornells, & Bachoud-Lévi (2007) hypothesized that if these authors were correct, ERP signatures elicited during rule learning would be different from those found in word segmentation studies. To test this hypothesis they created four artificial language streams formed by trisyllabic words following an A_iXC_i rule and separated by 25 ms pauses. These language streams were presented to participants for four minutes while ERPs were recorded. After the presentation

of each language, participants had to perform a two alternative forced choice test to evaluate both word and rule learning. In agreement with previous results (Cunillera et al., 2009, 2006; Sanders et al., 2002), a modulation in the N4 was already observed in the second minute of exposure to the languages, showing that in languages with non-adjacent dependencies, word learning can also be observed in brain activity after a short exposure. In addition, a P2 (a positive increase in amplitude at 200 ms after word onset) was detected in the third minute of exposure to the languages which correlated with rule learning scores from the behavioral test. When participants were divided into good and poor learners, only those who learned the rules displayed a progressive increase in the P2 amplitude through time (*see figure 3*). The authors attributed these findings to a shift of attention toward the regularities that had to be integrated. Indeed, several studies have shown that the P2 is modulated by perceptual learning and attention (Alain, Campeanu, & Tremblay, 2010; Guerreiro & Anguera, 2014; Reinke, He, Wang, & Alain, 2003), thus this change in attention would be endogenous since no external cues attracting participants' attention were presented during the study.

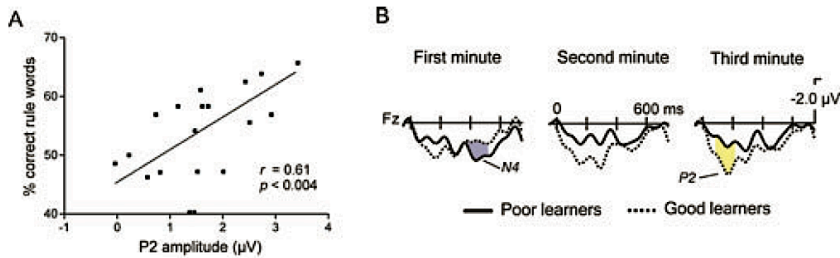


Figure 3. ERP results reported in De Diego-Balaguer et al. (2007). (A) Correlation between the mean amplitude of the P2 component at Fz in the third minute of learning (at the 120–220-ms time window) and performance on the rule-learning test. (B) ERP averages of the language conditions for each group at a frontal location (Fz), showing the evolution of the differences between groups over the learning period (first, second, and third minute). *Reproduced from De Diego-Balaguer et al., 2010.*

Mueller, Bahlmann, & Friederici (2008) investigated how the presence or absence of pauses in a speech stream formed by phrases (three bisyllabic words) following an A_iXC_i rule affect neural activity. Although they did not analyze brain activity during the learning phase, they found significant differences in the test phase between rule-based and random streams in the segmented and non-segmented conditions. Specifically, they found an increase in the P600 and a late anterior negativity (i.e., LAN), both previously associated with morphosyntactic and syntactic processes (Bahlmann, Gunter, & Friederici, 2006; Friederici, Steinhauer, & Pfeifer, 2002; Lelekov, Dominey, & Garcia-Larrea, 2000; Mueller, Hahne, Fujii, & Friederici, 2005). These findings indicate that pauses would trigger the extraction of syntactic regularities, as previous behavioral studies suggested (Endress et al., 2005; Peña et al., 2002), therefore supporting the hypothesis of the existence of two different mechanisms for word and rule learning.

Summing up, research with ERPs has greatly contributed to a better understanding of word segmentation and rule extraction. Several components have been attributed to different processes, suggesting the existence of two mechanisms. The appearance of both the N1 and the N4 indicate the segmentation of the speech stream. The N1 has been found to be more related to the detection of word onsets while the N4 reflects the construction of possible lexical candidates. Other components such as the P2, P600 and LAN were found during or after the process of rule extraction. The P2 has been interpreted as a shift of attention toward the cues that facilitate perceptual grouping, and the P600 and LAN have been related to the evaluation of sentences' correctness (i.e., syntactic judgment) after rule learning.

1.1.4.2. Neuroimaging studies of word and rule learning

Together with ERPs, neuroimaging studies have supported and complemented behavioural studies in which word and rule learning were investigated. Both structural and functional brain measures shed light on how and where in the brain these computations are processed, as well as the similarities they share when compared with natural language processing.

McNealy, Mazziotta, & Dapretto (2006) studied whether brain activity differed when three types of syllable streams were presented to participants: stressed (streams of four trisyllabic words stressed on the first syllable), unstressed (streams of four unstressed trisyllabic words) and random (streams of syllables placed in random order). More specifically, these researchers compared the increases in the BOLD signal between both stressed and unstressed versus random streams. From these comparisons they observed higher activation in the superior temporal gyrus, the transverse temporal gyrus and the inferior parietal lobe. In addition, in the test phase when brain activity elicited by words was compared with the activity elicited by part-words and non-words, higher activation in the posterior left inferior and middle frontal gyri were observed (*see figure 4*).

Similarly, Cunillera et al. (2009) studied brain activity when a group of participants faced a statistical language learning task. Their findings were in agreement with McNealy et al., (2006) since activation in the superior temporal gyrus and the premotor cortex (PMC) was found.

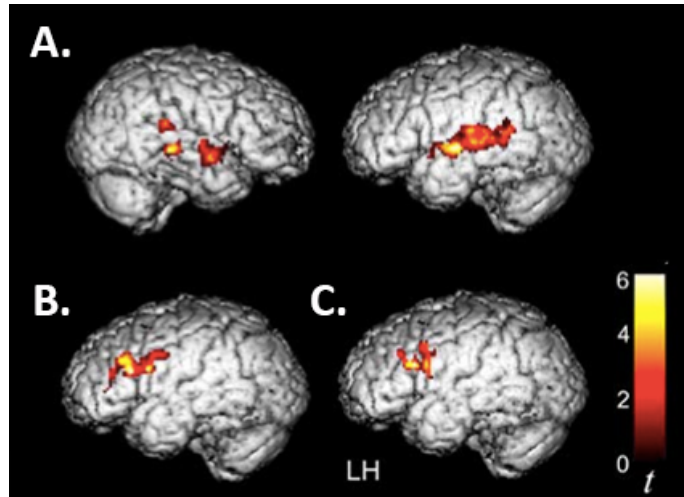


Figure 4. fMRI results reported in McNealy et al. (2006). (A) Increase of activity within the temporal cortex while listening the artificial language conditions presented in the study (i.e., stressed and unstressed). (B) Activity in the posterior left inferior and middle frontal gyri when listening to words compared with non-words in the unstress and (C) stress condition. *Adapted from McNealy et al., 2006.*

Results from both McNealy et al. (2006) and Cunillera et al. (2009) are in accordance with previous hypotheses on natural language learning and processing, proposing that word segmentation would be sustained by networks connecting frontal and temporal areas. The superior temporal gyrus would be essential for the mapping of the acoustic stimuli into phonological templates and/or sequences constructed from previous exposures to the stimuli (Warren et al., 2005). These templates would be forwarded to the PMC and they would be translated into articulatory gestures, thus allowing for them to be held in working memory (Friederici, 2009, 2012; Liao, Kronemer, Yau, Desmond, & Marvel, 2014; Saur et al., 2008). Indeed, the idea of feed-forward connections through the dorsal stream has been proposed as essential for both L1 and L2 learning

(Hickok & Poeppel, 2007; Poeppel & Hickok, 2004; Saur et al., 2008; Wong, Perrachione, & Parrish, 2007). Moreover, a recent study using DTI found that individual differences in the ability to learn new words were related to the structural properties of the left arcuate fasciculus (AF) (López-Barroso et al., 2013) – a WM structure that connects the temporal and frontal areas (Catani, Howard, Pajevic, & Jones, 2002; Catani, Jones, & Ffytche, 2005) (*see figure 5*). Even though participants in this study did not need to segment speech and, therefore, the use of statistical computations was not required, it shows that not only functional, but also structural properties of the dorsal pathway are relevant to explain individual differences in word learning (this topic will be extended in the next section).

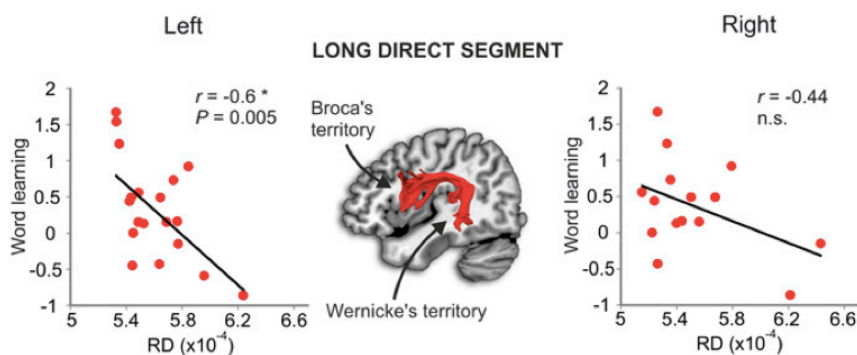


Figure 5. DTI results reported in López-Barroso et al. (2013). Link between the microstructure of the AF and word learning performance. The scatter plots show the correlations between word learning scores (d') and RD in the long AF for the left and right hemispheres. *Reproduced from López-Barroso et al., 2013.*

However, it is important to remark that other WM structures, such as the inferior fronto-occipital fasciculus (IFOF) or uncinate fasciculus (UF), also connect the temporal and frontal areas through

the ventral part of the brain (Catani et al., 2002; Catani & Thiebaut de Schotten, 2008). Different studies have shown the relevance of such structures in statistical learning tasks. In a case study from (Schapiro, Gregory, Landau, McCloskey, & Turk-Browne, 2014), a patient with complete bilateral hippocampal loss and broader middle temporal damage performed several statistical learning tests with different types of stimuli. Results reported a complete absence of this ability (compared with a control group), suggesting the implication of the damaged brain regions on the ability to extract temporal regularities. Importantly, both the IFOF and the UF have been linked to memory and attention (Catani & Thiebaut de Schotten, 2008; Doricchi, Thiebaut de Schotten, Tomaiuolo, & Bartolomeo, 2008; Gaffan & Wilson, 2008; Ross, 2008). Taking all of these data into consideration, although the AF seems to be of utmost importance for the acquisition of new lexical items, we cannot discard the possible implication of ventral WM structures on the computation of statistical regularities.

With regard to rule learning, we are not aware of any neuroimaging study investigating how and where grammar rules such as those described in Marcus et al. (1999) and/or morphological and syntactic rules based on non-adjacent dependencies (Peña et al., 2002) are processed in the brain. However, there is extensive literature on the neural correlates of artificial grammar learning (AGL) (Reber, 1967) – a paradigm in which participants are tested on the learning of grammatical patterns from strings of letters. A significant amount of AGL studies agree on the implication of

Broca's area on the processing of grammatical structures (fMRI: Forkstam, Hagoort, Fernandez, Ingvar, & Petersson, 2006; Opitz & Friederici, 2003, 2004; Petersson, 2004; TMS: de Vries et al., 2010; Uddén et al., 2008; DTI: Flöel, Vries, & Scholz, 2009). Likewise, researchers investigating natural language processing in healthy populations and patients with Broca's aphasia are in agreement with the AGL findings (de Diego Balaguer, Costa, Sebastián-Galles, Juncadella, & Caramazza, 2004; Shapiro, Moo, & Caramazza, 2012). Given that statistical learning and AGL require implicit learning of patterns containing temporal order information (Perruchet & Pacton, 2006), some authors have proposed that not only cortical, but subcortical structures (in particular the basal ganglia) would be involved in the acquisition of sequential patterns (for a review see Conway & Pisoni, 2008). The results of Karuza et al. (2013) are in agreement of such an assumption. These authors reported an increase of activation in the left inferior frontal gyrus and bilateral basal ganglia when forward versus backward speech containing statistical regularities were compared.

Taken together, neuroimaging studies on statistical learning have highlighted the role of the superior temporal and inferior parietal cortices on the analysis of the speech signal and the segmentation of phonological sequences (words) contained in speech. The left PMC may also play an important role on the maintenance in memory of the segmented words. Finally, studies with AGL suggested that mainly Broca's area is in charge of the processing and acquisition of grammar structures.

1.2. Factors contributing to individual differences in second language learning

People greatly differ in their ability to learn a new language. While some people can master an L2 in a short amount of time, others cannot achieve a good command of an L2 even if they were exposed to it in childhood or in a linguistic immersion environment.

A significant amount of research has focused on studying which aspects of the language system differentiate individuals the most during learning and which factors are responsible for these divergences (see for a review Sebastián-Gallés & Díaz, 2012). Two of the most important factors that determine language mastery are the age of acquisition and the amount of exposure (Ortega, 2014). However, not all aspects of a new language are equally affected by these two factors. For instance, mastering the speech sounds of an L2 is significantly more difficult if exposure to this language does not take place in childhood (Olson & Samuels, 1973), whereas the learning of both vocabulary appears relatively unaffected by age of acquisition (Jedynak, 2009). Moreover, it is evident that some individuals, despite of being exposed to an L2 for a short time and/or late in life, still have an excellent L2 performance (Hopp & Schmid, 2011). Such observations suggest the possession of an aptitude or talent, likely linked to a genetic advantage. It is important to note that current scientific views regarding the influence of genetic and environmental factors on human traits assume that neither genetics nor environment is uniquely

responsible for individual variation. Instead, most phenotypes would be explained in part by both factors, as well as their interaction. In the following sections both environmental and genetic factors contributing to individual differences in language learning are explained in more detail.

1.2.1. Environmental factors influencing language learning

Language development studies have provided significant insights about how certain aspects such as social interaction and input richness affect language learning (Hoff, 2006; Hoff, 2003). Along this line, Hart & Risley (1995, 1999) conducted a longitudinal study in which one-hour recordings were made by 42 families every month for two and a half years while they were interacting in a natural setting. The results from this study were interpreted to indicate that both the quantity and the quality of the language used by parents while they were speaking to their offspring were associated with the children's vocabulary learning. That is, they showed that vocabulary increased faster in those children who were receiving more input regardless of the socioeconomic status of their parents. However, follow-up studies claimed that it was not the quantity but the quality of the vocabulary as well as the non-verbal cues used by parents that really explained individual differences across infants (Cartmill et al., 2013; Hudon, Fennell, & Hoftyzer, 2013). As for grammar, studies revealed that vocabulary growth is associated with the subsequent evolution of grammar skills (Bates,

Friederici, Wulfeck, & Juarez, 1988; Marchman, Martinez-Sussmann, & Dale, 2004; Roberts, Burchinal, & Durham, 1999). However, in L2 learning, the pattern of acquisition of the lexicon versus grammar seems to be differentiated (Ullman, 2004, 2005).

A study carried out by Kuhl, Tsao, & Liu (2003) looked at how well infants learned a foreign language in two different types of situations: during real life social interaction or during a television and audio session. Results showed that those children exposed to the television and the audio sessions scored significantly lower than those who were exposed to the life session. Hence, the study suggested that social interaction when learning a new language would be of significant help.

As mentioned previously, the age at which an L2 is acquired is an important factor in L2 mastery. The existence of specific periods in development, called critical (or sensitive) periods, where brain structures are more susceptible to external inputs (Hensch, 2005) supports such assumption. In the case of L2 learning, it is important to consider that not all aspects of language learning are equally influenced by age of acquisition; each aspect is likely to have a different maturational timing (Hickok & Poeppel, 2007; Huttenlocher & Dabholkar, 1997; Huttenlocher, 1999). For example, auditory perception is mainly processed in the temporal areas with very early maturing timing (Huttenlocher & Dabholkar, 1997; Pujol et al., 2006), hence proficiency in L2 phoneme perception would be significantly more difficult after very early childhood. On the other hand, vocabulary learning could be

achieved later in life since its processing is more extended across brain areas which reach maturation in childhood (and even adolescence) (Huttenlocher & Dabholkar, 1997; Leroy et al., 2011). In summary, since different aspects of language are processed in brain regions with different maturational times, critical periods may vary significantly across them.

Results from the studies presented above demonstrate that social interaction and the quality of the input are important factors for L1 learning. In L2 acquisition, the age of acquisition is crucial and varies across language aspects (e.g., phoneme perception or grammar). However, other aspects must also be taken into consideration. For instance, the time of exposure, learning intention (Flege & Liu, 2000), L2 immersion (Linck, Kroll, & Sunderman, 2009) as well as non-language related factors (e.g., perinatal problems (Gerner & Baron, 2014), nutrition (Fretham, Carlson, & Georgieff, 2011), musical experience (Slevc & Miyake, 2006; Wong et al., 2007)) may also be important determinants of L2 mastery.

1.2.2. Twin and genetic studies related to language

The complexity of both genome and language make research in this field extremely challenging. Different methodological approaches have been taken to uncover the determinants of language learning variability. Most of the research has been focused on either

language development or language impairment, leaving L2 learning virtually unexplored. However, there is no evidence that the same genetic profiles which cause language problems and/or the causes of variability observed in developmental studies are also responsible for individual differences in L2 learning.

The following section will briefly introduce two methods used to investigate the influence of genetic and environmental factors on specific traits (e.g., physical traits such as weight or height, psychiatric traits such as schizophrenia, cognitive traits such as intelligence and/or behavioral traits such as extroversion). In addition, twin studies focused on language disorders in which speech processing was one of the main problems are presented. Finally, the section focuses on findings related to normal language development, including the few studies targeting L2 learning.

1.2.2.1. The relevance of twin studies

Why are twins ideal for the study of the origin of individual difference? Two types of twins are found in the population: monozygotic twins (MZT) and dizygotic twins (DZT). The main difference between them is that MZT share almost 100% of their genetic material while DZT share only 50%, on average. Assuming that both types of twins have similar environmental backgrounds, the following reasoning applies: if a specific trait is more similar between MZT than DZT, it would suggest that genetics play a role in this trait. But, if MZT are not more similar to each other than

DZT, it suggests that the heritability for this trait is low and that environmental factors play a greater role. Based on these principles, different experimental approaches allow inference of the influence of genetics and environment on a given trait. Introduced here are only the most relevant studies necessary to understand the methodologies used in this dissertation (for further reviews on this topic see Boomsma, Busjahn, & Peltonen, 2002 and van Dongen, Slagboom, Draisma, Martin, & Boomsma, 2012).

In 1875 Francis Galton published what is considered the first study of twins titled “*The history of twins, as a criterion of the relative powers of nature and nurture*”. However, it was not until 1914 that Poll, followed by Siemens (1924) published the first studies using the ‘*classic*’ twin design. This design has been applied to estimate the effect of three parameters on a phenotype:

1. The genetic effect (heritability) (A).
2. The shared environment effect (C).
3. The unique environment effect (E).

As the acronym for each parameter indicates, this model is called “ACE” and it measures the proportion of the variance in a trait that is heritable as well as the proportions which are due to shared environmental or unshared environment.

Heritability is defined as the proportion of the phenotypic variation (V_P) that is due to variation in genetic values (V_G) (Falconer & Mackay, 1996). Individuals in a population (except MZT) vary in

the genotypes they have at the loci. The combined effect of all loci is the genotypic value. Therefore, heritability can be formulated, in its simplest form, as $H^2 = V_G/V_P$.

As mentioned earlier, heritability is most frequently estimated by comparing resemblances between MZT and DZT. Thus, Falconer developed the following formula: $H^2 = 2(r_{MZT} - r_{DZT})$, where a heritability index is considered high when the correlation between MZT is twice that of DZT. However, two other factors must be considered: first, the environmental factors that twin pairs share (common environmental factors), and second, factors that are not shared (specific environmental factors) due to different life experiences (e.g., injuries, studies, social relationships). (*see figure 6 for some examples of traits in which the ACE model has been applied*).



Figure 6. Examples of results from classical twin analysis. Percentage of variances explained by genetic factors (purple), by shared environmental factors (green), by unique environmental influences (beige) and by differences in age (blue). The number in the brackets gives the modal age of the sample in years. Phenotypes include cardiovascular risk factors, and personality and cognitive traits. *Reproduced from Boomsma et al, 2002.*

In spite of its relevance, this model does not consider several factors that may interfere when measuring heritability. That is, in order to define heritability as a main effect of genetic variation, the model ignores the possible presence of Gene X Environment interactions, the transmission of epigenetic information through the germline or the maternal RNA molecules to the fetus, the biochemical state of the gametes during conception and/or the entrance of nutrients, bacteria or antibodies into the embryo, etc. (Bettegowda & Smith, 2007; Boulinier & Staszewski, 2008; Chong & Whitelaw, 2004; Grindstaff, Brodie, & Ketterson, 2003; Hasselquist & Nilsson, 2009; Rassoulzadegan et al., 2006; Shorter & Lindquist, 2005).

To this end, another powerful method used to understand the effect of environment on a particular trait or disease is the *discordant twins design*. Since MZT are genetically matched and usually have the same environmental background, any discordance between them should be explained by experiences that are unique to each twin (i.e., non-shared environmental factors). The discordant MZT design is used to study the probable causal pathway between exposure to non-shared environmental experiences and a specific characteristic. It is based on the identification of discordant twin pairs with respect to the outcome of interest (i.e., each twin's score belongs to one of the extremes of a continuum) and comparing them on possible predictors (e.g., De Moor, Boomsma, Stubbe, Willemsen, & de Geus, 2008). Neuroimaging studies have used this design to explore the causes of hand and language laterality (Badzakova-Trajkov, Häberling, & Corballis, 2010; Geschwind,

Miller, DeCarli, & Carmelli, 2002; Häberling, Badzakova-Trajkov, & Corballis, 2013) or to identify environmentally-shaped brain regions involved in a given trait or pathology (de Geus et al., 2007) (*see figure 7*).

A variation of this method, the *MZT differences design*, has been also used to assess the influence of non-shared environment on a phenotype (Pike, Reiss, Hetherington, & Plomin, 1996; Plomin, DeFries, Knopik, & Neiderhiser, 2013; Turkheimer & Waldron, 2000). In this case instead of selecting discordant twins, the researchers measure MZT dissimilarities in a continuum. That is, the within-pair differences for the environmental factor, as well as for the outcome variable (i.e., the trait to be studied), are calculated by subtracting the score of one twin from the score of the co-twin. Correlations and regression analysis are employed to evaluate the relationship between differences in the environmental factor and differences in the outcome. This method has been used mainly in behavioral developmental studies. For example, researchers studied how parenting affects certain aspects of children's behavior (Asbury, Dunn, Pike, & Plomin, 2003; Caspi et al., 2004; Deater-Deckard et al., 2001; Yamagata et al., 2013), how life events influence sleep quality and chronotype (Barclay, Eley, Buysse, Maughan, & Gregory, 2012; Barclay, Eley, Parsons, Willis, & Gregory, 2013), or even how parenting, life events and social relationships impact depression (Liang & Eley, 2005). The advantage of this design compared to the discordant twin design is that there is no need to find extremely discordant MZT, as intra-pair

similarity is expected to vary continuously between pairs of twins.

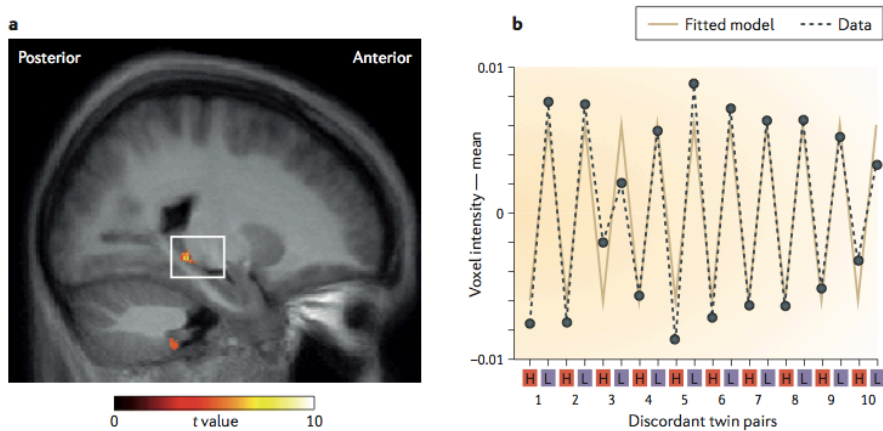


Figure 7. Results reported in de Geus et al. (2007). (a) Image shows the left parahippocampal area, where a significant volume reduction was found in the high-risk twin compared with the low-risk co-twin from MZT who were discordant for anxious depression. Colors represent the effect size (t value from paired t test) of the comparison of grey matter volume between twins. (b) Graph shows the relative responses (individual voxel intensity minus mean voxel intensity in all twins) of ten discordant twin pairs at the most significant voxel in the left parahippocampal area ('H' indicates the twin with high risk of anxious depression and 'L' indicates the low-risk co-twin). *Reproduced from de Geus et al., 2007.*

1.2.2.2. Twin studies in language disorders

A major issue in the study of language impairment is the extensive heterogeneity of existing disorders and their comorbidities. However, specific language impairment (SLI) is of particular interest given that children that suffer it shows dysfunctional statistical learning (Evans, Saffran, & Robe-Torres, 2009), impaired ability to extract non-adjacent dependencies (Hsu, Tomblin, & Christiansen, 2014), impaired word learning (Nash & Donaldson,

2005) and grammar processing (D. V Bishop, Chan, Adams, Hartley, & Weir, 2000). Moreover, SLI is one of the language disorders in which genetic and environmental factors have been investigated most.

One of the characteristics of SLI is the limitation that children exhibit in the phonological short-term memory system, an ability needed for word learning (Gathercole, 2006) and word segmentation (Cunillera et al., 2006; Lopez-Barroso et al., 2011; McNealy et al., 2006). Bishop et al. (1999) studied the heritability of phonological working memory in a group of 7-12 year-old twins diagnosed with SLI. The test they used was a non-word repetition task in which a list of nonsense words was presented to children and they were asked to repeat them. Results showed a significant influence of genetics on this ability. Interestingly, subsequent molecular genetic studies supported these findings by showing a region on chromosome 16 with significant linkage to non-word repetition performance (SLI Consortium, 2002, 2004).

Children with SLI exhibit a deficit in producing and comprehending morphological inflections, such as the past tense and plural (Leonard & Eyer, 1996). It has been suggested that these errors are caused by SLI children's inability to learn and formulate rules, that is, they can learn to produce some inflections correctly by memorizing them but cannot generate them in novel words (Gopnik & Goad, 1997). Importantly, several studies have shown a significant genetic influence in children with this particular

limitation (Bishop et al., 1999; Critten, Connelly, Dockrell, & Walter, 2014). Rice, Haney, & Wexler (1998) examined the family history of SLI children whose main deficit was grammar learning. They found more language impairments in first-degree relatives of SLI children compared to non-affected children, indicating that verb inflection marking in SLI may be under genetic control.

Taking into account that both phonological working memory and grammar processing in SLI display high heritability indices, Bishop, Adams, & Norbury (2006) questioned whether these two abilities were under the same genetic influence. To answer this question, they studied the ability of 173 six year-old twin pairs to perform both a non-word repetition task and a verb inflection task. Afterwards heritability was estimated. Results replicated a previous study (Bishop, 1999) by revealing high heritability for both abilities (61% for phonological working memory and 74% for grammar). However, analyses also showed that the phenotypic and etiological overlap between them was minimal. Thus, authors suggested that the capacity for retaining strings of unfamiliar speech sounds for short periods of time and the capacity to perform grammar-like computations have distinct genetic origins.

1.2.2.3. Language twin studies in healthy populations

Twin studies investigating language in normal populations have mostly centered on language development (Stromswold, 2001). One

example is the research carried out by Dionne et al. (2003) in which the genetic and environmental relationship between vocabulary and grammar was investigated in a sample of more than 2000 twins between 2 and 3 years of age. Specifically, they administered the MacArthur Communicative Development Inventory, which is a standardized parent reporting system used to assess children's lexical growth. They applied a multivariate genetic model that revealed modest heritability for both vocabulary (i.e., 21% at 2 years of age and 10% at 3 years of age) and grammar (i.e., 42% at 2 years of age and 34% at 3 years of age), indicating that grammar was more influenced by genetic factors than was vocabulary. Both Dale, Dionne, Eley, & Plomin (2000) and Dionne et al. (2003) found a high genetic overlap between these two variables, leading the authors to suggest a common origin. However, these results are opposed to Bishop et al's (1999) in children with SLI. Thus, this disagreement might indicate that the genetic causes leading to abnormal development may not be the same as those leading to individual differences in normal language development, a hypothesis that still needs investigation.

Following a procedure similar to that of Dionne et al. (2003), other twin studies investigated the involvement of genetics in vocabulary and grammar learning in older children, obtaining different heritability indices. Dale, Harlaar, Hayiou-Thomas, & Plomin (2010) reported a genetic influence of 30% in both vocabulary and listening grammar in 12-year-old children; in Byrne et al. (2006) the genetic sources explained 32% of the variance for vocabulary and

29% for grammar in kindergarten children; finally, in Kovas & Hayiou-Thomas (2005) heritability accounted for 52% of the variance for vocabulary and 21% for grammar in 4.5 year-old children. A meta-analysis of five studies of vocabulary in 3- to 12-year-old children (Fischer, 1973; Foch & Plomin, 1980; Mather & Black, 1983; Segal, 1984; Thompson, Detterman, & Plomin, 1991) in which 330 MZT and 237 DZT were included, reported that genetic factors accounted for 53% of the variance (Hoekstra, Bartels, van Leeuwen, & Boomsma, 2009). Thus, taken all together, most of the studies included in this section reported a moderate influence of genetics in both vocabulary and grammar in children between 3 and 12 years of age, although there was some variability across studies.

Van Soelen et al. (2009) went further in the study of vocabulary and investigated heritability of the capacity to memorize and forget a list of words, depending on the length of the time period between the learning and the test. They presented a list of words aurally to 10 and 14 year-old twins and their siblings. Participants were asked to recall the words immediately after the presentation, as well as after a period of 20-30 minutes. Results showed higher correlation for both learning and forgetting speed in MZT than in DZT and siblings, with heritability estimated at 43% for learning and 20-30% for forgetting. This study is particularly interesting because, although the task used was not directly related to word segmentation, it is a good indicator of verbal working memory, an ability needed to accomplish word learning (Lopez-Barroso et al.,

2011).

One of the few twin studies investigating L2 acquisition asked whether training to conjugate English verbs from present to past tense had an effect on brain activation in 13-year-old twins (Sakai, Miura, Narafu, & Muraishi, 2004). They evaluated seven MZT and two DZT pairs before and after 2 months of classroom training. Results showed high correlations between twins in both groups, both in performance when they were examined in the classroom as well as inside the scanner. Moreover, the increase in activation in the inferior frontal gyrus during the task also significantly correlated between twin pairs. Although, in this study heritability was not measured, it suggested that both genetic and environment significantly determine our ability to learn verb morphology in an L2. Dale, Harlaar, Haworth, et al. (2010) evaluated the ability of 14-year-old twins to write, listen, read and speak in both L1 and L2. Researchers observed higher heritability in L2 (67%) compared to L1 (46%). Moreover, they discovered a high overlap of genetic factors between L2 and L1, suggesting that the proportion of individual differences explained by genetics is shared for both L1 and L2.

Taken together, the results of twin studies in normal populations indicate that genetic factors explain approximately 30-40% of variance in vocabulary and grammar learning, as well as verbal working memory. Although genetic factors affecting general cognitive abilities influence linguistic performance (Hohnen &

Stevenson, 1999), data support the hypothesis that language is partly the result of innate predispositions. Furthermore, they suggest that those genetic factors affecting L1 learning are also involved in L2 learning.

1.3. The present research: Goals, procedures and hypotheses

The main goal of the present work is to have a better understanding about the origin of individual differences observed when language is learned. As it has been presented in this introduction, two fundamental mechanisms in language acquisition have been suggested: word and rule learning (Marcus et al., 1999; Peña et al., 2002; Saffran, Newport, & Aslin, 1996). Importantly, studies investigating these two processes reported individual differences in brain structure and activity (Abla et al., 2008; De Diego Balaguer et al., 2007; López-Barroso et al., 2013; McNealy et al., 2006; Sanders et al., 2002). For instance, it has been shown that some specific ERP waves are present in good learners: the N1 and N4 waves (for word learning) (Alba et al., 2005; Sanders et al., 2002) and the P2 wave (for rule learning) (De Diego Balaguer et al., 2007). Similarly, significant correlations have been reported between performance in word learning and both functional activity in premotor areas (McNealy et al., 2006) and structures such as the AF (López-Barroso et al., 2013). Hence, it is known that being good or poor on learning words and rules correlates with individual differences in

our brain structure and function, however it is still unclear the origin of the variation observed across population. Here we address, for the first time, the extent to which genetic and environmental sources explain individual differences in word and rule learning.

Two studies are included in this dissertation. In the first study we took advantage of the fine-grained temporal information that electrophysiological measures provide to investigate how brain activity elicited during word and rule learning is defined by genetic and environmental factors. In order to do that we followed the experimental procedure developed by De Diego Balaguer et al. (2007) who evaluated participants on both word and rule learning while ERPs were recorded. These authors constructed four speech streams formed by words following AXC rules (Peña et al., 2002). Words were separated by 25 millisecond pauses to facilitate rule learning. Our interest in this study was the significant correlation reported between the amplitude of the P2 and the rule learning performance. Importantly, it was the first time that individual differences in neural activity were related to the ability of learning rules. Hence, De Diego Balaguer et al.'s (2007) procedure was optimal to our goals as it elicited individual differences and also in a single experiment it allowed the study of word and rule learning. In order to estimate heritability, we compared MZT and DZT twins. By measuring the intra-pair similarity and by applying Falconer's heritability formula, the contribution of genetics on the resulting word and rule learning neural activity was calculated. If we rely on the data provided by previous heritability studies (Byrne et al.,

2006; Kovas et al., 2005) we would hypothesize a moderate contribution of genetic factors on both types of learning, while predicting a slightly higher contribution for word acquisition. However, we must take into account that these studies tested infants in their first language, therefore, it is possible that genetic and environmental factors would have a different effect in adults' language learning.

The second study aimed at discovering how environmental pressures in language WM structures are associated with individual differences in statistical learning. López-Barroso et al. (2013), by using artificial language, investigated whether the AF was the structure responsible for the auditory-motor integration previously observed during word learning. Interestingly, they found significant correlation between word learning and the microstructural properties of the AF (long segment). Hence, based on this study we centered our research on the AF as well as the IFOF and the UF, two tracts also connecting frontal and temporal areas. There are three main differences in this second study compared to the first one. In the first place, here we were interested in studying individual differences in the ability to segment words, hence the speech stream used here had neither pauses nor rules. In the second place, we focused on brain structure instead of brain activity. We used DTI to obtain information about WM properties in those structures previously shown (or hypothesized) to be involved in language processing (i.e., AF, IFOF and UF). And finally, instead of investigating heritability, we focused on those non-shared

environmental factors that may explain variability in word learning and its WM substrates. A MZT differences design was the optimal approach to investigate such non-shared environmental effects. Hence, the correlation between intra-pair differences on word learning scores and intra-pair differences on WM was measured. By doing so we could know what structures affected by non-shared environmental pressures were related to individual differences in word segmentation. Since AF has been considered the main route connecting Wernike's area with preomotor and Broca's areas (Catani et al., 2002, 2005) we hypothesized that mainly environmental pressures on this tract would explain individual differences in statistical learning.

2. EXPERIMENTAL SECTION I

Manuscript 1

Basora A., Jin Y., Burgaleta M. and Sebastián-Gallés N.
(Submitted) **Low heritability of individual differences
in word and rule learning: A twin study of Event-
Related potentials.** *PlosOne*

**Low heritability of individual differences in word and rule
learning: A twin study of Event-Related Potentials**

Anna Basora^{1*}, Yu Jin¹, Miguel Burgaleta¹ and

Núria Sebastián-Gallés¹

1. Center for Brain and Cognition, Department of Technology,
Universitat Pompeu Fabra, Roc Boronat 138, 08018 Barcelona,
Spain.

** Corresponding author:*

Anna Basora
Center for Brain and Cognition
Department of Technology
Universitat Pompeu Fabra
C/ Roc Boronat, 138
08018 Barcelona, Spain
Voice: +34 935 422 362
E-mail:anna.basora@upf.edu

Abstract

In this research we aimed at exploring the genetic and environmental contributions to the ability for extracting words and rules from speech. Event-related potentials were obtained from monozygotic and dizygotic twins while they were listening to four artificial language streams. A wavelet analysis was carried out and cross-correlations were used to measure the similarity in brain activity for each pair of siblings. Results showed that in both word and rule learning conditions, monozygotic twins were more similar than dizygotic twins in both the first and fourth minute of exposure to the languages. This indicates that brain activity is influenced by genetic factors from an early learning stage. However, the low heritability index obtained would suggest that environment is of utmost importance.

Keywords: twins, language learning, event-related potentials, wavelet, heritability.

1. Introduction

People differ in their ability to learn a second language (L2). Such individual variability can arise at different levels, e.g. when learning to discriminate non-native speech sounds or when trying to learn new words (W) and grammatical or syntactic rules. These differences have been attested in both natural settings [1,2] and artificial languages [3,4]. However, the underlying sources of individual differences in language learning are not well understood. Studies investigating L2 acquisition point out the importance of environmental factors such as the amount of exposure to the new language or its age of acquisition [5]. Moreover, it has been shown that other factors, such as motivations or social aspects like parents interaction are fundamental to acquire a language [6]. Nevertheless, the role played by genetic influences has received, comparatively, little attention. Most studies on the genetic contribution to language performance have explored the origin of (developmental) speech and language disorders [7,8,9]. However, it has been acknowledged that only a small fraction of these disorders can be explained solely by genetic abnormalities [8]. Furthermore, it is an open question whether the same genetic causes leading to abnormal development are also at the basis of individual differences in L2 learning in individuals with a normal acquisition of the first language. Along this vein, research on the genetic origins of language in healthy populations has provided estimates of the extent to which different language functions are heritable, although with highly heterogeneous results ranging from very low to very high

heritability indices depending on the targeted ability and the specific measure used [10].

Crucially, previous studies on the genetic sources of language functions have paid little to no attention to L2 learning. Furthermore, the available research often relies on paper-and-pencil questionnaires that are unable to provide fine-grained information about language processing and its time course. Our goal is to understand the basic mechanisms underlying differences in L2 learning by investigating their genetic and environmental sources. Because the age of acquisition and amount of exposure are two of the most important factors in determining L2 proficiency [11], one methodological challenge is to properly control participants' previous exposure to the L2. One way of sorting out this problem is to use artificial languages, thus ensuring that all participants start from scratch. This approach allows us to estimate an individual's capacity to compute statistical regularities present in speech (i.e., statistical learning), a domain-general ability that has been suggested to partly account for individual differences in language learning [4,12]. From this perspective, words and language rules can be understood as a combination of elements (e.g., speech sounds) that display a high probability of appearing concatenated (transitional probability (TP)) in natural speech. By experimentally manipulating transitional probabilities in artificial speech streams, some studies have demonstrated that human adults, infants [13,14,15,16] and even animals [17,18] are able to capture the statistical information present in speech input, and it has been

shown that this ability is supported in humans by brain structures tightly linked to language functions (i.e., the left arcuate fasciculus; [19]). At the functional level, event related potential studies have provided fine-grained evidence about the neural substrates of statistical language learning processes. More specifically, the N4 event-related potential (ERP) component has been suggested as an index for the construction of a linguistic trace in the lexical repertoire [3,20–22], whereas the P2 component is sensitive to the learning of morphological rules [3]. Currently, however, the degree to which these language learning ERP indices can be accounted for by genetic or environmental factors remains unexplored.

Here, a sample of monozygotic (MZT) and dizygotic twins (DZT) was selected, and a classical twin design was applied in order to obtain estimates of the relevance of genetic and environmental factors on ERP-based measures of word and rule learning. The classical twin design has been frequently used to measure the contribution of genetics and environment on specific phenotypes ranging from physical or medical traits to behavioral or social ones [10] . It is based on the comparison between MZT, who share all segregating genetic material, and DZT, who share, on average, half of their segregating genetic material. Hence, if genetics plays a significant role on a given phenotype, MZT will show higher similarity than DZT.

Likewise, ERPs have been successfully applied to investigate the genetic and environmental influences on language-related brain

activity [23,24]. In our study, MZT and DZT performed a previously validated word and rule learning task [3] while ERPs were recorded in order to measure their brain responses during the learning processes. We presented to the participants four artificial language streams (4 min each) formed by trisyllabic words following an A_iXC_i rule (see Methods) separated by short pauses. Afterwards, twins were faced with a two alternative forced choice test to assess word and rule learning. ERPs analyses were focused on the two components studied by [3]: the N4 for word learning and the P2 for rule learning. Firstly, to ensure that we replicated previous findings, we tested the whole sample for an increase in the amplitude of both components (indexing a learning effect) from the first to the fourth minute of exposure to the artificial language streams. Secondly, from the EEG we performed a continuous wavelet transformation analysis, allowing us to explore not only amplitude and latency information but also how the elicited signal was distributed through different frequency bands, thus providing a complete picture of the learning process. From the full wavelet plot we selected the time windows of the components we were interested in and we measured the cross-correlation between the two component wavelets from each pair of twins [25,26]. Finally, the average cross-correlations for MZT group were compared with the average cross-correlation of DZT. Since previous findings have suggested that individual differences in word and rule learning are not solely an effect of environmental exposure, we expected to find higher correlations for MZT than for DZT on the wavelet time windows corresponding to P2 and N4. However, since it is the first

time that a study of such characteristics has been performed, no specific hypotheses can be formulated with regard to the magnitude of the differences between groups when N4 and P2 wavelet time windows are compared.

2. Materials and Methods

2.1. Participants

60 healthy monozygotic twins (30 pairs) (46 females, mean age = 21.51, SD= 4.32) took part in this study. Also, a sample of 52 healthy dizygotic twins (32 females, mean age = 21.33, SD= 3.09) was included. Each pair of DZT was of the same sex. None of them reported any language or hearing deficits or had specific music knowledge.

2.2. Ethics statement

The experiment was approved by the local ethics committee of the Universitat Pompeu Fabra and it was in compliance with the Code of Ethics of the World Medical Association (Declaration of Helsinki). Written consent was obtained from each participant prior to the experiment. Since two of the participants were underage (17 years old), written consent from their legal guardians was also required. At the end of the experiment underage participants received a T-shirt and the rest of participants received a monetary compensation.

2.3. Stimuli and Procedure

The procedure was identical to the one used by [3], but new stimuli were created. Both members of each twin pair came to the laboratory on the same time slot (morning or afternoon) and they were tested one after the other. All participants sat in a comfortable armchair in a sound-attenuated and electrically shielded room and all the auditory stimuli were delivered binaurally through headphones at an intensity of 70 dB.

Four languages formed by nine trisyllabic words were created. That is, words were built following a rule which defined that the initial syllable determined their ending regardless of the middle one (**bidome**, **bifime**, **bigume**) [27]. There were 3 different frames and 3 different middle syllables, which resulted in a total of 9 different words per language (*see Table 1*). The items were synthesized with MBROLA, a speech synthesizer software [28] based on the concatenation of diphones at 16 kHz from the Spanish male database (es2). Word duration was 696 ms and they were separated by 25 ms pauses in order to facilitate the extraction of the structural information. Words were pseudo-randomly concatenated with the restriction that the consecutive repetition of the same word was not allowed. The TP between the first and the middle syllable and between this one and the last syllable was 0.5. The TP between the non-adjacent syllables (first and third) was equal to 1 within words and it was 0.5 between the last syllable of any word and the first syllable of the following one.

The experiment was divided into three different phases: learning, violation and recognition. Four artificial languages streams formed the learning phase, each one with a duration of 4 minutes (336 words). Participants were instructed to pay attention to the language they would listen to because they would have to recognize words from that language later on.

In the violation phase the same language streams as the previous phase were presented but in this case, non-words (NW) and rule-words (RW) were inserted pseudo-randomly (four to ten words between each test item). Non-words were formed as words but in the inverse order (i.e. word: **bidome**, non-word: **medobi**). Rule-words were formed following the same rule as words but the middle syllable was replaced with one belonging to another word (i.e. word: **bidome**, rule-word: **bitame**). Each item was presented twice for each language (18 NW and 18 RW), which represents the 5.3% of the stimuli. Participants were unaware of the changes applied in this new phase. They were told to continue listening to the languages as in the previous phase.

The recognition phase consisted in the presentation of a two-alternative forced choice test. It evaluated whether participants had acquired the words and learnt the rules from the previous language streams. The test items were presented in pairs with a separation of 704 ms. Participants had to choose between words and non-words in order to test word acquisition, and between rule-words and non-words in order to test rule learning (i.e., both rule- and non-words

were the same as the ones presented in the violation phase). Each test item appeared twice, leading to 18 W, 18 RW and 144 NW presentations per language stream. Participants had to listen to the two alternative items and respond by pressing a right or left button after a cross appeared on the screen.

Table 1. Stimuli used for each of the artificial languages.

	Structure	Middle Syllables	Word	Non-word	Rule-word
Language 1	pa__su ne__ta bi__me	do, fi, gu	padosu	sudopa	pabisu
Language 2	be__pi du__ro so__ku	ma, li, te	bemapi	pimabe	bedupi
Language 3	ti__pu ga__mo ra__de	go, le, ki	tigopu	pugoti	tidepu
Language 4	pe__di ri__na ba__lo	se, mu, fo	pesedi	disepe	penadi

For each language nine words could be made combining each middle syllable with each structure. Non-words were formed using the same words' syllables but placed in inverse order. Rule-words were constructed using the same three structures, and the middle syllable belonged to another word. In the words, non-words and rule-words columns are presented one example for each language.

2.4. Electrophysiological recording and preprocessing

The EEG was continuously recorded with an online filter (band pass of 0.1 to 100 Hz) and digitized (A/D rate= 512 Hz). Tin electrodes mounted in an electro cap (Electro-Cap International) were used to measure EEG and electroculogram (EOG) activity. Recordings were made at 32 scalp locations according to 10/20 system and two additional electrodes placed at the left (LM) and right (RM) mastoids. Blinks and horizontal eye movements were measured with electrodes attached to the infra-orbital ridge and on the outer canthus of the right eye. The reference for the EEG and the EOG was located on the tip of the nose. Electrode impedances were kept below 5 kOhms.

ERP analysis was focused on the learning phase since it was the phase where individual differences were observed in [3]. Moreover, the paradigm was originally designed to compare ERPs across groups, thus the number of trials included in both the violation and the recognition phase were not enough to obtain a robust signal for an inter-subject comparison. This phase was segmented into four blocks of one minute duration each. Analysis was time-locked to the onset of the stimulus and mean amplitude of the time windows of interest were averaged for each minute of exposition. Eye movements were corrected following the Gratton and Coles Method [29] implemented in the Brain Vision Analyzer Software package (v. 2.0; Brain Products GmbH, Munich, Germany). Epochs with EEG exceeding either $\pm 100 \mu\text{V}$ at any channel, activity lower than $0.5 \mu\text{V}$, or voltage step/ sampling greater than $50 \mu\text{V}$ within

intervals of 200 ms were automatically rejected offline. Epochs were of 700 ms, including 100 ms of pre-stimulus baseline. For individual segments the baseline was corrected and a lineal DC detrend procedure was applied. ERPs were digitally band-pass filtered between 0.1 and 30 Hz (slope of 12 dB/oct). Only participants with more than the 70% of free artifact epochs for each language were included in further analyses. Four MZT pairs and two DZT pairs were excluded.

2.5. Learning analysis

2.5.1. Behavioral analysis

The percentage of accuracy for W and RW was calculated and compared between groups. One subject from the MZ group was not included in the analysis due to software malfunction and data loss.

First, a one-tailed *t*-test against 50 level was performed in order to know whether twins scored above chance level. Second, two two-tailed *t*-tests were computed comparing MZT and DZT in both W and RW conditions.

2.5.2. ERP analysis

As a first step, in order to replicate potential learning effects previously found [3] we compared the brain activity in the first and fourth minute of exposure to the unknown languages. In this step we also verified whether MZT and DZT responded to the stimuli in the same way. To this end and following previous studies [3,20,21] we selected two time windows to be analyzed: a 150-250 ms time

window (P2 component) and a 300-500 ms time window (N4 component) on the Fz electrode, since [3] found inter-individual significant effects at this specific location.

We performed a linear mixed effects regression analysis [30] with mean amplitude as the dependent variable using the implemented *nlme* package [30] in R software [31]. Twin Type (MZT and DZT) and Minute (Min1 and Min4) were defined as fixed effects (with an interaction term). As random effects, we had intercepts for Siblings, which was common to a pair of twins, and for Monozygotic, which was shared by a pair of MZT but not by a pair of DZT [32].

2.6. Analysis of similarity

A wavelet analysis was performed in order to compare the brain activity from one twin with his/her co-twin. Importantly, this approach allows for a more complete image of the signal and it provides more information for inter-subject comparison. For each participant, the wavelet power spectrum was computed at the same electrode location (Fz) we chose in the time-amplitude analysis. A continuous wavelet transformation was performed following the Mexican hat method [33].

Epochs of 4000 ms, including 1000 ms pre-stimulus appearance, 600 ms stimulus presentation, and 2400 ms post-stimulus presentation were used in the wavelet analysis [34]. The selected frequency range was set between 0.5 and 30 Hz with 50 frequency

steps. The long epoch used in the analysis was cut 100 ms before the stimuli and 600 ms after to avoid edge effects during wavelet transformation (*see Figure 1*).

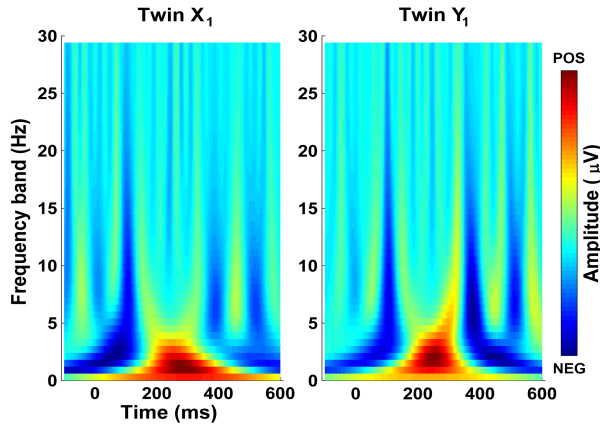


Figure 1. Example of two wavelet maps from a pair of twins. In each plot the time (ms), frequency (Hz) and relative amplitude (μV) are represented.

The intra-pair comparisons were performed for a 150-250 ms window (P2 component), as well as for a 350-450 ms window (N4 component). Note that the window for N4 was 50 ms smaller compared to the previous analysis (i.e., ERP amplitude analysis), in order to have the same data points in the two windows which was required to fairly compare the two windows in the intra-pair analysis. Moreover, after visual inspection of the signal for each couple of twins, we observed that many of them presented small time lags in the peak of the components, which led to underestimations of wavelet similarity in the windows of interest. In order to solve this problem we opted for a cross-correlation analysis (CCA). CCA provides information on the similarity between two signals and the time lag between them. The formula applied was the following:

$$X_{f,t} = \begin{pmatrix} a_{1,1} & \cdots & a_{1,t} \\ \vdots & \ddots & \vdots \\ a_{f,1} & \cdots & a_{f,t} \end{pmatrix} \quad (1)$$

$$Y_{f,t} = \begin{pmatrix} a_{1,1} & \cdots & a_{1,t} \\ \vdots & \ddots & \vdots \\ a_{f,1} & \cdots & a_{f,t} \end{pmatrix} \quad (2)$$

$$r_{xy}(f,t) = \frac{\sum_{f,t} X(f,t) * Y(f,t)}{\sqrt{(\sum_{f,t} X(f,t) * X(f,t)) * (\sum_{f,t} Y(f,t) * Y(f,t))}} \quad (3)$$

By applying this approach, the overlap of the activity patterns within the selected window was measured and a maximum value indicated when two data vectors are "most alike". Visual inspection was performed to ensure that this method yielded appropriate results. By applying this approach, slight displacements of the selected windows in the time domain were allowed in order to increase the overlap of the activity patterns. The maximum intra-pair cross-correlation value for each window was submitted to a repeated measures ANOVA, comparing the twins in the two minute and time windows. Specifically, the factors Minute and Window were considered as within-subject factors and the two groups of participants were taken as between-subject factor (Twin Type factor). Additionally, we computed a heritability index applying Falconer's formula [35]:

$$H^2 = 2(r_{mzt} - r_{dzt})$$

Heritability shows the proportion of a trait variance that is explained by the genetic variance.

3. Results

3.1. Learning Results

3.1.1. Behavioral results

Prior to the genetic (similarity) analyses, we analyzed behavioral and ERP data to ensure that participants showed a learning pattern. The mean score in the word-learning test was 65.8% for MZT and 65.4% for DZT, and both groups performed above chance level (MZT: $t_{(50)}=12.13$, $P<0.001$; DZT: $t_{(47)}=10.41$, $P<0.001$). In the rule-learning test, the performance was at chance level for both groups (MZT: 50.9%; DZT: 51.4%), consistently with other reports where brain activity signals for statistical learning were identified in absence of behavioral evidence [20,36]. Two-tailed t -tests were applied to compare MZT and DZT in both word-learning and rule-learning conditions. No significant differences were found between groups (Words: $t_{(97)}=0.216$, $p=0.830$; Rule-words: $t_{(97)}=0.283$, $p=0.778$) (see Figure 2).

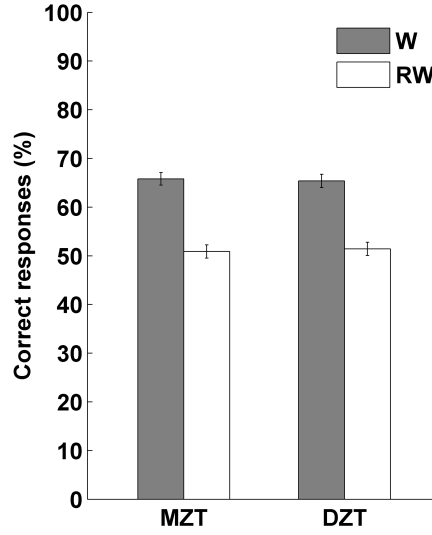


Figure 2. Percentage of accuracy from the Recognition phase. Correct responses for MZT and DZT are shown in the W and RW learning tests. Error bars indicate standard error (SE).

3.1.2. ERP results

First of all, mean activity values at the Fz electrode were taken as the dependent variable and the two time windows were submitted to a linear mixed effects regression analysis. As expected, for the P2 we observed a significant effect of Minute ($\beta=0.276$, S.E.=0.079, $t_{(126)}=3.456$, $p<0.001$) but not of Twin Type ($\beta=0.118$, S.E.=0.093, $t_{(48)}=1.263$, $p=0.213$), nor the interaction term ($\beta=0.034$, S.E.=0.126, $t_{(126)}=0.304$, $p=0.782$). For the N4, the same pattern of results was observed. The effect Minute showed differences between the first and fourth block at a trend level ($\beta=0.168$, S.E.=0.091, $t_{(126)}=1.845$, $p=0.067$) but no difference between groups or Group by Minute interaction was observed (All $p>0.5$). The main effect of Minute in P2 and N4 components replicated previous findings [3] and showed an effect of learning after four minutes of exposure to the language stream (see Figure 3).

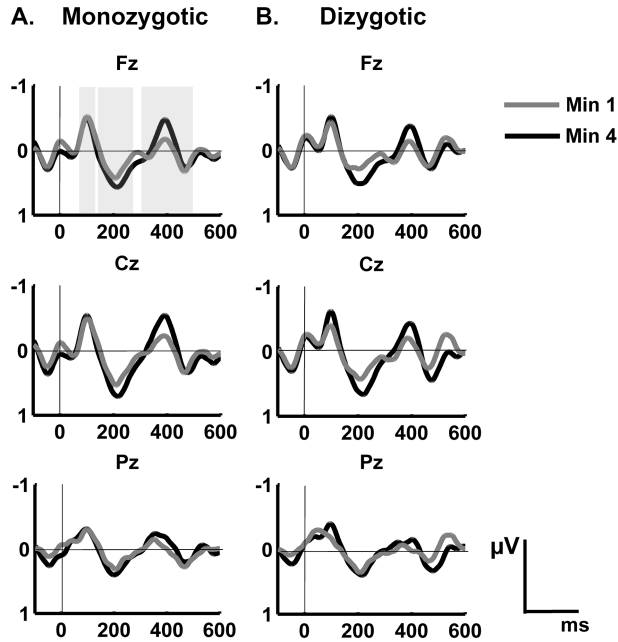


Figure 3. ERPs grand average at Fz, Cz and Pz electrodes for the Learning phase pooled across the four languages. Gray areas from the first graph indicate the time windows selected for N1, P2 and N4 analysis. **A.** ERP averages for MZT twins comparing the first (dotted line) and the fourth (full line) minute of exposition to the languages. **B.** ERP averages for DZT twins comparing the first and the fourth minute of exposition to the languages. In both cases an increase of the P2 and N4 was observed from the first to the fourth minute.

Additionally, visual inspection of the ERPs revealed a N1 component not observed by [3], but reported in other studies using continuous speech segmentation [20,21]. Therefore, this component was also included in our analyses taking a window of 60 ms (70-130 ms) centered in its maximum peak [21]. Results showed no significant main effects nor interaction between the factor Twin Type and Minute (All $p > 0.5$).

3.2. Similarity Results

A 2x2x2 repeated measures ANOVA was performed comparing wavelet intra-pair similarity for the two groups (Twin Type factor) in the two minutes (Minute factor) and the two time windows (Window factor). Three main effects were found to be significant (Minute: $F_{(1,50)}=12.421$, $p=0.001$; Window: $F_{(1,50)}=10.213$, $p<0.002$; Twin Type: $F_{(1,50)}=7.577$, $p=0.008$) but no significant interactions were observed. (see Figure 4).

To estimate heritability we applied Falconer's formula [34] to the mean cross-correlation for MZT and DZT in the fourth minute for both time windows. The value obtained for the P2 window was $H^2=0.163$ and $H^2=0.189$ for the N4 window.

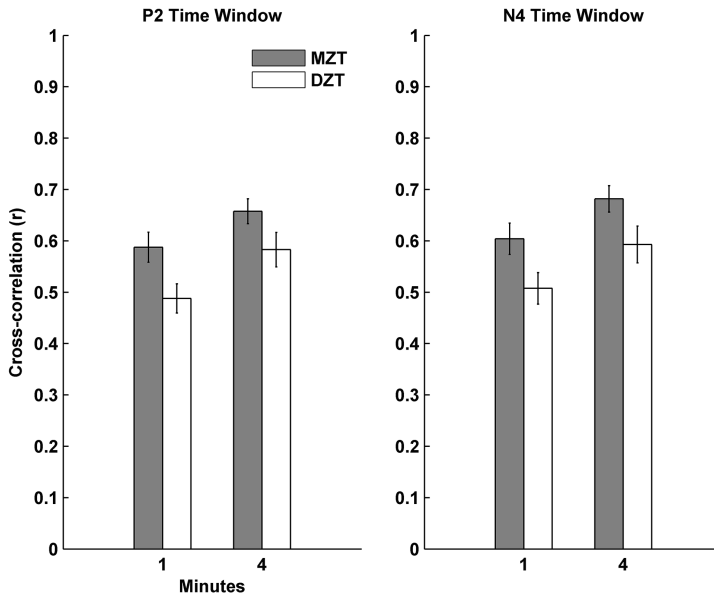


Figure 4. Average intra-pair maximum cross-correlation for MZT and DZT. MZT and DZT cross-correlations are shown from the wavelet time window corresponding to the rule learning (P2) and word learning process (N4) in the first (1) and fourth (4) minute of exposition to the language.

4. Discussion

The main goal of this study was to explore, for the first time, the influence of genetics and environment in word and rule learning processes in an artificial language. We selected a sample of MZT and DZT and they performed an artificial language learning task while ERPs were recorded. Consistently with previous reports, participants showed a learning ERP pattern [3,20,21,37] reflected in a significant increase of the N4 and P2 amplitudes from the first to the fourth minute of exposure to the language. Importantly, similarity analyses (comparing wavelet cross-correlations in the N4 and P2 time windows between genetic groups) revealed low heritability indices, thus suggesting an important role of environmental factors in word and rule learning.

Our main result was driven by a greater similarity for MZT compared the DZT in both the first and the fourth minute of exposure, as well as in both time windows. This indicates that genetic factors were, as expected, of relevance when explaining individual differences in language learning, even at early stages of the learning process. Nevertheless, the magnitude of the difference between MZT similarity and DZT similarity (summarized in the H^2 index) underscores the relevance of environmental factors.

Few twin studies have emphasized the importance of environmental influences in word learning. [38] have studied vocabulary in 5 year-old twins. Similar to our results, they reported a null contribution of genetics and they emphasized the importance of cultural and socio-

economic status in word learning [39]. Also, [40] investigated receptive language skills in 12-year-old twins (N=4,892) and the results showed that heritability for vocabulary and grammar is low and that non-shared environment is the most predominant factor, explaining 58% of the variance for vocabulary and 54% for grammar. Unfortunately, twin studies on L2 learning are scarce. One example is [41], in which 13-year-old twins were tested on verb identification and verb matching (in L1 and L2) during a scan session. Authors found significant correlations between siblings at the behavioural and brain activity levels and concluded that mechanisms underlying L2 acquisition are highly determined by both genetic and environmental factors.

It is noteworthy that the duration of the exposure to the artificial language had an impact on the similarity estimates, so that an increase in wavelet cross-correlations was observed between the first and fourth minutes for both groups. We suggest that this increase in cross-correlation is caused by the consolidation of the learning process. Specifically, at the beginning of the exposure to the languages, participants did not know the words they were listening to, and this uncertainty should be reflected on a more variable brain activity across participants. However, after four minutes, most participants already would have learned the words, and brain activity would then be expected to become more stable across subjects [34,41–43], therefore contributing to an overall convergence of the wavelet signals.

Interestingly, we also observed differences in similarity between time windows. That is, the cross-correlation across groups and minutes was higher for N4 (word learning) than for P2 (rule learning). These differences are consistent with the time course of both word and rule learning. Although only two minutes of exposure to a continuous language stream can suffice to segment and learn words [14], it is likely that after four minutes of exposure the brain activity in the N4 time window will become more stable. However, this might not be the case for rule learning. Data from infants [44] and adults [3] have shown that more time might be necessary for rule extraction. Hence, after four minutes, it is possible that some of the participants did not implicitly acquire the rules yet, and the activity in P2 time window might be more unstable across subjects. Indeed, the results from the behavioral test seem to corroborate this argument. Hence, the temporal pattern for the learning process would account for the higher cross-correlation in the N4 time window compared to the P2.

In conclusion, this is the first time that individual differences in word and rule learning using ERPs have been addressed from a nature-versus-nurture perspective. Moreover, to our knowledge, no previous twin study measured the contribution of genetic and environmental factors at the beginning and the end of the word learning process, thus providing online genetic/environmental variations not accessible from the conventional questionnaire testing approach. Our results show that after just the first minute of exposure to the language, MZT are significantly more similar than

DZT, and this difference persists after four minutes of exposure. These results might indicate that brain activity during word and rule learning is influenced by genetic factors even when learning is not yet consolidated. However, the low heritability index obtained would suggest that the environment may be an essential factor for word and rule extraction.

References

1. Dörnyei Z (2006) Individual differences in second language acquisition. *AILA review* 19(1): 42–68.
2. Robinson P (1997) Individual Differences and the Fundamental Similarity of Implicit and Explicit Adult Second Language Learning. *Lang Learn* 47(1): 45–99.
3. De Diego Balaguer R, Toro JM, Rodriguez-Fornells A, Bachoud-Lévi A-C (2007) Different neurophysiological mechanisms underlying word and rule extraction from speech. *PLoS One* 2(11): e1175.
4. Misyak JB, Christiansen MH (2012) Statistical Learning and Language: An Individual Differences Study. *Lang Learn* 62(1): 302–331.
5. Birdsong D (1999) *Second Language Acquisition and the Critical Period Hypothesis*. E-Library T& F, editor Mahwah: Lawrence Erlbaum Associates, Inc.
6. Kuhl PK, Tsao F-M, Liu H-M (2003) Foreign-language experience in infancy: effects of short-term exposure and social interaction on phonetic learning. *Proc Natl Acad Sci U S A* 100(15): 9096–9101.
7. Newbury DF, Monaco a P (2010) Genetic advances in the study of speech and language disorders. *Neuron* 68(2): 309–320.
8. Kang C, Drayna D (2011) Genetics of speech and language disorders 1. *Annu Rev Genomics Hum Genet* 12: 145–164.
9. Graham S, Fisher SE (2013) Decoding the genetics of speech and language. *Curr Opin Neurobiol* 23(1): 43–51.
10. Stromswold K (2001) The Heritability of Language: A Review and Metaanalysis of Twin, Adoption, and Linkage Studies. *Language* 77(4): 647–723.

11. Flege JE, Bohn O-S, Jang S (1997) Effects of experience on non-native speakers' production and perception of English vowels. *J Phon* 25(4): 437–470.
12. Misyak JB, Christiansen MH, Bruce Tomblin J (2010) Sequential Expectations: The Role of Prediction-Based Learning in Language. *Top Cogn Sci* 2(1): 138–153.
13. Endress AD, Bonatti LL (2007) Rapid learning of syllable classes from a perceptually continuous speech stream. *Cognition* 105(2): 247–299.
14. Saffran JR, Aslin RN, Newport EL (1996) Statistical learning by 8-month-old infants. *Science* 274(5294): 1926–1928.
15. Saffran JR, Newport EL, Aslin RN (1996) Word segmentation: The role of distributional cues. *J Mem Lang* 62(3): 606–621.
16. Lany J, Saffran JR (2013) Statistical Learning Mechanisms in Infancy. *Comprehensive Developmental Neuroscience: Neural Circuit Development and Function in the Healthy and Diseased Brain*. Amsterdam: Elsevier Inc., Vol. 3. pp. 231–248.
17. Hauser MD, Newport EL, Aslin RN (2001) Segmentation of the speech stream in a non-human primate: statistical learning in cotton-top tamarins. *Cognition* 78(3): B53–64.
18. Toro JM, Trobalon JB, Sebastián-Gallés N (2005) Effects of Backward Speech and Speaker Variability in Language Discrimination by Rats. *J Exp Psychol Anim Behav Process* 31(1): 95–100.
19. López-Barroso D, Catani M, Ripollés P, Dell'Acqua F, Rodríguez-Fornells A, et al. (2013) Word learning is mediated by the left arcuate fasciculus. *Proc Natl Acad Sci U S A* 110(32): 13168–13173.

20. Sanders LD, Newport EL, Neville HJ (2002) Segmenting nonsense: an event-related potential index of perceived onsets in continuous speech. *Nat Neurosci* 5(7): 700–703.
21. Cunillera T, Toro JM, Sebastián-Gallés N, Rodríguez-Fornells A (2006) The effects of stress and statistical cues on continuous speech segmentation: an event-related brain potential study. *Brain Res* 1123(1): 168–178.
22. Mueller JL, Bahlmann J, Friederici AD (2008) The role of pause cues in language learning: the emergence of event-related potentials related to sequence processing. *J Cogn Neurosci* 20(5): 892–905.
23. Anokhin AP, Heath AC, Myers E (2006) Genetic and environmental influences on frontal EEG asymmetry: a twin study. *Biol Psychol* 71(3): 289–295.
24. Almasy L, Porjesz B, Blangero J, Goate a, Edenberg HJ, et al. (2001) Genetics of event-related brain potentials in response to a semantic priming paradigm in families with a history of alcoholism. *Am J Hum Genet* 68(1): 128–135.
25. Young JP, Lader MH, Fenton GW (1972) A twin study of the genetic influences on the electroencephalogram. *J Med Genet* 9(1): 13–16.
26. Hornickel J, Lin D, Kraus N (2013) Speech-evoked auditory brainstem responses reflect familial and cognitive influences. *Dev Sci* 16(1): 101–110.
27. Peña M, Bonatti LL, Nespor M, Mehler J (2002) Signal-driven computations in speech processing. *Science* 298(5593): 604–607.
28. Dutoit I, Pagel C, Pierret N, Baraille E (1996) The MBROLA project = Towards a set of high quality speech synthesizers free of use for non commercial purposes. 3: 1393–1396.

29. Gratton G, Coles MG, Donchin E (1983) A new method for off-line removal of ocular artifact. *Electroencephalogr Clin Neurophysiol* 55(4): 468–484.
30. Pinheiro JC, Bates DM (2000) *Mixed-Effects Models in S and S-PLUS*. 1st ed. Springer.
31. R Core Team (2013) *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria.
32. Visscher PM (2004) Power of the Classical Twin Design Revisited. *Twin Res* 7(5): 48–54.
33. Shark L-K, Yu C (2006) Design of matched wavelets based on generalized Mexican-hat function. *Signal Processing* 86(7): 1451–1469.
34. De Diego-Balaguer R, Fuentemilla L, Rodriguez-Fornells A (2011) Brain dynamics sustaining rapid rule extraction from speech. *J Cogn Neurosci* 23(10): 3105–3120.
35. Falconer D, Mackay T (1996) *Heritability. Introduction to quantitative genetics*. 4th Ed. San Francisco: Benjamin Cummings. pp. 160–183.
36. McNealy K, Mazziotta JC, Dapretto M (2006) Cracking the language code: neural mechanisms underlying speech parsing. *J Neurosci* 26(29): 7629–7639.
37. Cunillera T, Càmarà E, Toro JM, Marco-Pallares J, Sebastián-Galles N, et al. (2009) Time course and functional neuroanatomy of speech segmentation in adults. *Neuroimage* 48(3): 541–553.
38. Kavé G, Shalmon M, Knafo A (2013) Environmental contributions to preschoolers' semantic fluency. *Dev Sci* 16(1): 124–135.

39. Hoff E (2003) The Specificity of Environmental Influence: Socioeconomic Status Affects Early Vocabulary Development Via Maternal Speech. *Child Dev* 74(5): 1368–1378.
40. Dale PS, Harlaar N, Hayiou-Thomas ME, Plomin R (2010) The etiology of diverse receptive language skills at 12 years. *J Speech Lang Hear Res* 53(1): 982–992.
41. Sakai KL, Miura K, Narafu N, Muraishi Y (2004) Correlated functional changes of the prefrontal cortex in twins induced by classroom education of second language. *Cereb Cortex* 14(11): 1233–1239.
42. Hasson U, Malach R, Heeger DJ (2010) Reliability of cortical activity during natural stimulation. *Trends Cogn Sci* 14(1): 40–48.
43. Abrams DA, Ryali S, Chen T, Chordia P, Khouzam A, et al. (2013) Inter-subject synchronization of brain responses during natural music listening. *Eur J Neurosci* 37(9): 1458–1469.
44. Gómez R, Maye J (2005) The Developmental Trajectory of Nonadjacent Dependency Learning. *Infancy* 7(2): 183–206.

3. EXPERIMENTAL SECTION II

Manuscript 2

Basora A., Burgaleta M., Sebastián-Gallés N. (Submitted)

**Environmental contributions to the white matter
structures associated with word learning.** *Proceedings
of the National Academy of Sciences*

Classification: BIOLOGICAL SCIENCE. Neuroscience

Title: Environmental contributions to the white matter structures associated with word learning

Short title: Environment and language learning

Authors: Anna Basora¹, Miguel Burgaleta^{1*} and Nuria Sebastian-Galles¹

Affiliation: 1. Center for Brain and Cognition, Department of Technology, Universitat Pompeu Fabra, Roc Boronat 138, 08018 Barcelona, Spain.

** Corresponding author:*

Miguel Burgaleta, PhD
Center for Brain and Cognition
Department of Technology
Universitat Pompeu Fabra
C/ Roc Boronat, 138
08018 Barcelona, Spain
Voice: +34 935 422 362
E-mail: miguel.burgaleta@upf.edu

Abstract

It is well known that not everyone is equally good at learning a second language, but the origin of these individual differences is not well understood. Recently it has been shown that the integrity of the arcuate fasciculus partly accounts for individual differences in word learning [López-Barroso D, Catani M, Ripollés P, Dell'Acqua F, Rodríguez-Fornells A, & de Diego-Balaguer R (2013). Word learning is mediated by the left arcuate fasciculus. *Proc Natl Acad Sci USA* 110(32):13168–73]. Here we asked to what extent environmental factors contribute to the association between word learning and its white matter substrates. With this purpose, white matter diffusivity measures in the arcuate fasciculus and other potentially relevant connections (inferior fronto-occipital fasciculus and uncinate fasciculus) were acquired from a sample of monozygotic twins. Also, participants underwent a statistical learning task that yielded estimates of word learning ability. Finally, a twin differences design was applied, so that the association between pairwise (between-siblings) differences in word learning as well as in white matter integrity was tested. Significant results were found for arcuate and inferior fronto-occipital fasciculi, thus suggesting that environmental pressures drive the association between word learning and the integrity of its white matter underpinnings.

Keywords: word learning, arcuate fasciculus, inferior fronto-occipital fasciculus, diffusion tensor imaging, monozygotic twins

Significance Statement

Why do people differ in their ability to learn a new language? While this crucial question remains unanswered, the available literature suggests that experience may be a main determinant of individual differences in the development of language brain networks. Here we show that monozygotic twins (MZT) that differ in their ability to learn new words also present differences in the integrity of their language connections. Because MZT share almost 100% of their genetic material, these differences must stem from exposure to different environments. Our results support that environmental pressures shape white matter connections essential for language learning.

1. Introduction

People differ in their ability to acquire a second language (L2), including how good they are at learning new words. In the recent years brain imaging studies have provided important knowledge about the neural substrates underlying this ability (e.g., 1–3). Functional magnetic resonance imaging (fMRI) studies have revealed that a left-lateralized frontotemporal network is active when individuals identify new words in the speech stream (4, 5, 2). At the structural level, white matter (WM) connections supporting such functional network are also crucial for word learning (2). More specifically, it has been shown that microstructural properties of the left arcuate fasciculus (AF), assessed by means of diffusion tensor imaging (DTI), are related to performance in a word learning task.

Nevertheless, although the available evidence highlights that the transmission of information between temporal and frontal areas through the AF is fundamental for word learning, it is unclear to what extent individual differences in this fronto-temporal network plastically arise across development and adulthood due to environmental factors. The importance of environmental factors such as social interaction (6, 7), age of acquisition (8–12) and amount of exposure (13, 14) is well documented by studies of L2 learning. Twin studies quantifying the relevance of genetics and environment have yielded highly heterogeneous estimates (15–18), while paying little to no attention to L2 learning. Notably, language-related WM heritability has been shown to be of moderate magnitude (30-40%) and decreasing with age (19), thus suggesting

that environmental pressures may be crucial for shaping language learning abilities.

Based on recent findings in which AF has been associated to individual differences in word learning (2), here we wanted to make a step forward by studying the environmental effect on these individual differences. More specifically, we investigated the role of environmental factors on the development of those WM structures relevant for word learning by applying a *twin differences design*, that has been previously used to assess the influence of non-shared environment in a particular trait (20–22). The rationale of this approach assumes that, because monozygotic twins (MZT) are characterized by sharing virtually a hundred percent of their DNA, any discordance in their phenotype must arise from differential exposure to environmental elements. This method has been successfully applied to uncover how parenting styles affect different aspects of children behavior (23–26), how life events influence sleep quality (27) and chronotype (28) or the impact of parenting styles, life events and social relationships on depression (29). Moreover, in neuroimaging studies this method has been applied in order to identify the environmentally-shaped brain regions involved in certain pathologies such as depression, anxiety or obsessive-compulsive disorder (30–33). We scanned a sample of 40 MZT who underwent a diffusion tensor imaging (DTI) acquisition who also performed a statistical word learning task. In order to establish what environmentally-shaped WM structures are related to the ability to learn new words, the intra-pair differences in performance were

correlated with the intra-pair differences in WM integrity from three potentially relevant WM connections: the AF (2), the inferior fronto-occipital fasciculus (IFOF) (34, 35) and the uncinate fasciculus (UF) (36). We hypothesized that if WM substrates of word segmentation are sensitive to non-shared environmental pressures, then there should be a positive correlation between these two measures.

2. Results

After the acquisition of diffusivity measures and the presentation of two artificial language streams, a 2 and a 4 minute-long one, twins performed two word-learning tests. The mean scores for each test were 56.9% (SD=17.12) and 69.74% (SD=19.41) respectively. Both conditions were performed above chance level (2 minutes: $t_{(38)}=2.49, p=0.017$; 4 minutes: $t_{(38)}=6.27, p<0.001$) (see *Figure 1*).

The difference between twins in absolute value divided by the sum was measured for the means of the behavioral tests and the four DTI measures (fractional anisotropy (FA), axial diffusivity (AD), radial diffusivity (RD) and mean diffusivity (MD)) for left and right AF, IFOF and UF. Two outlier pairs (more than 2 SD from the mean) in the AF (FA measure) were excluded. Significant correlations were found between the 2-minute language learning scores and the left AF (FA: $r=0.741, p=0.001$; AD: $r=0.496, p=0.031$) as well as between the 4-minute language learning scores and the left IFOF (FA: $r=0.489, p=0.034$; AD: $r=0.460, p=0.048$; RD: $r=0.522,$

$p=0.022$; MD: $r=0.543$, $p=0.016$). No significant correlations were observed in the right hemisphere or in the UF (*see Figure 2*).

In addition, partial correlations between structural and behavioral measures were calculated controlling for the Siblings factor, which was common to each pair of twins. Results did not show any significant correlation for any of the tracts.

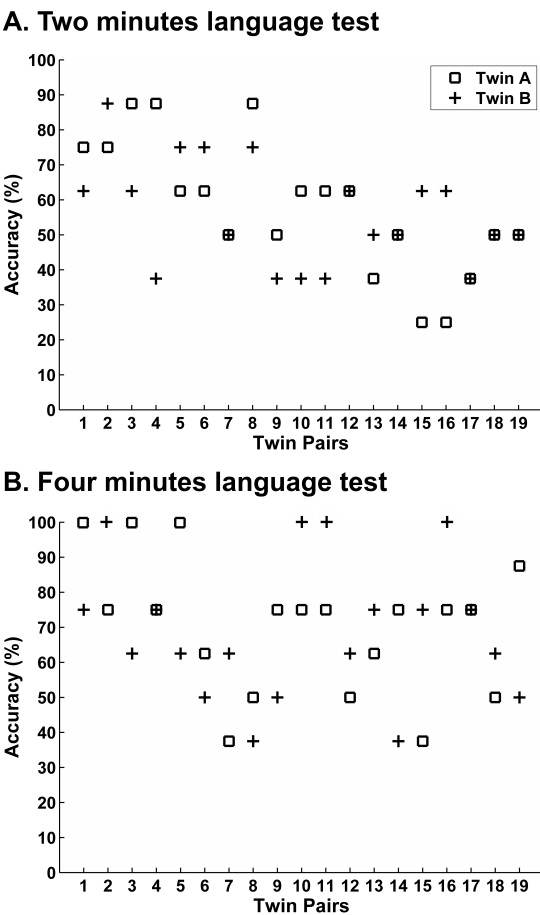
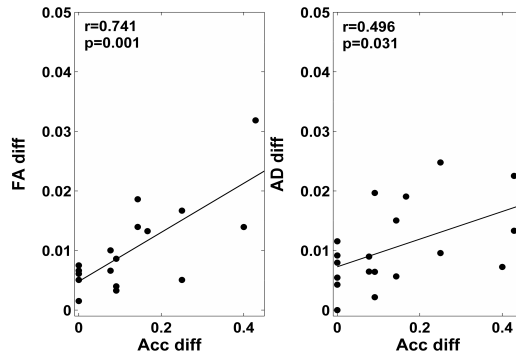


Figure 1. Percentage of accuracy for each pair of twins in the two (A) and four (B) minutes language tests.

A. Left arcuate fasciculus



B. Left inferior fronto-occipital fasciculus

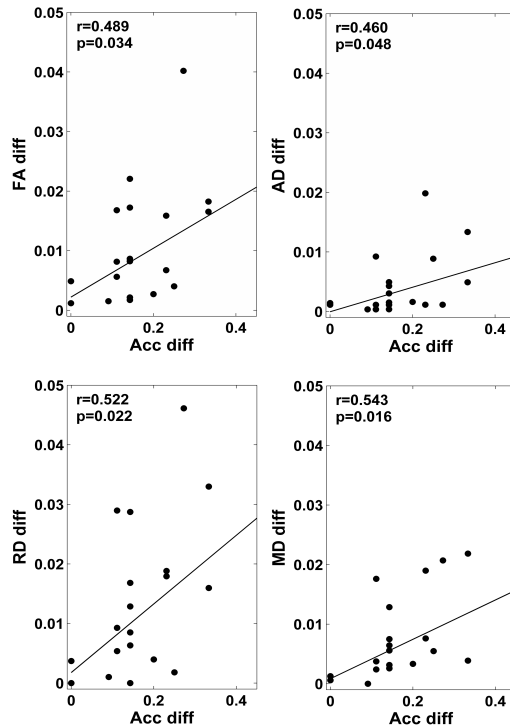


Figure 2. Relation between the twin intra-pair differences of the diffusivity measures and the percentage of word learning accuracy. **A.** Significant correlations between intra-pair difference of 2 minutes language learning scores and differences in FA (N=17) and AD (N=19) for the left arcuate fasciculus. **B.** Significant correlations between intra-pair differences of 4 minutes language learning scores and differences in FA, AD, RD and MD (all N=19) for the left inferior fronto-occipital fasciculus.

3. Discussion

The main goal of this study was to investigate whether the association between the left AF and statistical language learning is affected by environmental factors. MZT were tested in a statistical learning task and the intra-pair difference in performance was correlated with intra-pair differences in WM integrity from three potentially relevant WM connections: AF, IFOF and UF. Results confirmed the relevance of the AF, but also of the IFOF, for statistical learning and showed, for the first time, that twins who differed in their learning scores also differed in their AF and IFOF microstructure. Since MZT generally have the same genetic architecture and, in our sample, had been raised in the same household, differences between them would likely emerge due to non-shared environmental factors. Hence, differences in the ability to segment and learn new words would be related to experience-dependent structural properties of the left AF and IFOF. Therefore, here we went further with respect to previous research that reported an association between individual differences in statistical word learning and microstructural variability of the left AF (2).

Specific roles for IFOF and AF

Interestingly, AF was sensitive to early language stages (after a 2 minute exposure) whereas the effect in the IFOF was detected in the latest exposure stage (minute 4). This distinction fits well with the hypothesized stages of statistical language learning, namely, that we first segment the speech streams and secondly, we store words in memory once they are segmented. AF appears to be especially

involved in a first word-learning step where the phonological units of the speech stream are analyzed and their regularities are extracted. Indeed, several studies have shown that AF is responsible for the phonological processing during both comprehension and production (37, 38). The feedback and forward AF connections between auditory and motor areas would allow the maintenance of information in working memory when participants are listening to the speech stream (5, 39). More specifically, the new chain of speech sounds would activate the speech production system to create new articulatory sequences, which would become stronger representations of the new words (40). The IFOF, on the other hand, would be involved in the second step required for word learning, namely the consolidation of the auditory representation of the new words (41) and the search for conceptual associations (42). IFOF is part of the language ventral pathway, that is responsible for mapping the acoustic speech input onto meaningful items with different levels of complexity (i.e., segments, syllabic structure, phonological word forms, grammatical features and semantic information) (43). Given that it connects frontal and occipital areas, IFOF has been mainly associated to reading skills (44, 45) and semantic processing (34, 35). However, recent studies have reported its involvement in the maintenance of verbal information (46–48) and its retrieval (49), a possibility that is further supported by IFOF's prefrontal pattern of connectivity – essential for verbal working memory (50, 51). Hence, it is likely that IFOF plays a role in the long-term storage of the new words in the lexicon. This fits well with the observation that only two minutes are needed to

segment a speech stream (52), and thus subsequent exposure to the language may help to reach a better memorization and performance, therefore increasing accuracy.

Environmental factors in white matter development

The relevance of the environment highlighted in our study is consistent with the available evidence regarding WM sensitivity to environmental pressures. WM is characterized by its significant development during childhood and adolescence (53–55). A considerable amount of studies have reported the promotion of axonal myelination based on learning experience, thus supporting the effect of environment on the structural formation of the brain (56–58). Twin studies also reflected the importance of environment in WM (19, 59) and how it becomes more prominent as a function of age. For instance, heritability of WM has been shown to be lower in adults compared to newborns (60). Moreover, a recent multi-site study combined datasets from several DTI twin studies including 2248 twins from different races/ethnicities and ages (9-85 years old) (61). Data were pooled to obtain a robust FA heritability estimate. Results from the meta- and mega-genetic analysis showed that heritability estimates for both superior longitudinal fasciculus and IFOF left-sided tracts were lower than the right-sided estimates. Thus, these findings are consistent with the fact that we observed environmental effects on the left AF and IFOF.

Interpretation of diffusivity measures and possible involvement of other WM structures

In our study we investigated four diffusivity measures (i.e. FA, AD, RD, MD), that have been associated to different WM features (62). FA has been shown to be sensitive to fiber coherence, myelin thickness and axonal integrity; AD would describe the direction and organization of the fibers in the track; RD is associated to myelination, axon diameter and fiber density; and finally, MD is considered either a complementary or alternative measure to FA (63). The results showed that both FA and AD differences correlated with word learning differences in the AF. All four measures significantly correlated with differences in word learning scores at the IFOF. Hence, on the one hand, the correlations found in the AF suggest that the environmental pressures that link statistical word learning with AF morphology take place at the level of myelination and fiber orientation. On the other hand, results for the IFOF indicate that myelination and a range of different morphological features of the fibers may be affected by language experience.

Nevertheless, microstructural properties of other WM pathways may also be involved in statistical learning –especially given that it is a very basic mechanism present also in infants and other species, such as monkeys and rats (64, 65), and AF and IFOF are not well developed in these individuals (14, 66). In particular, since AF is not well developed in newborns, it has been suggested that the ventral pathway could be an alternative route for language

processing (67). Other studies have reported activation in the basal ganglia during statistical language learning (4), and in sequence learning (68–71). We cannot discard that environmental modifications of other WM structures (e.g., those connected to the basal ganglia) also contribute to the individual differences observed in our study.

Conclusion

This is the first time that it has been shown how environmental factors are relevant to explain the link between word learning and its WM substrates. Our results provide evidence of the impact of experience on language-related WM and more specifically, they show that environmental changes in WM structures may be responsible for the individual differences previously observed in statistical learning and word learning.

4. Methods

4.1. Participants

40 healthy MZT (20 pairs) (32 females, mean age 22.70, S.D.= 4.01) participated in this study. All of them were right handed and no one reported any language or hearing problem and no neurological or psychiatric disorders. The experiment was approved by the local ethical committee of the Universitat Pompeu Fabra and it was in compliance with the Code of Ethics of the World Medical Association (Declaration of Helsinki). Written consent was obtained from each participant prior to the experiment. All participants were paid at the end of the experiment for their participation.

4.2. Stimuli and procedure

The experiment consisted of two parts, a behavioral test and an MRI session. While one of the twins was in the scanner the other one was performing the behavioral test. When both twins finished the first session, they switched tasks.

4.3. Behavioral testing

In the behavioral test the participants listened two artificial language streams. Languages were composed by four tri-syllabic words pseudo-randomly concatenated with the restriction that the consecutive repetition of the same word was not allowed. Words were 696 ms long and they were synthesized with MBROLA, a speech synthesizer software (72) based on the concatenation of diphones at 16 kHz from the Spanish male database (es2) (<http://tcts.fpms.ac.be/synthesis/>). All the syllables used in the two languages were different and they were created following previous studies (52, 73–75). The transitional probabilities between syllables were 1 and 0.33 for within and between words, respectively. Languages were presented one after the other and the order of presentation was the same for each sibling but counterbalanced across twin pairs. One of the languages was two minutes long and the other one was four minutes long without breaks or pauses between words. A different duration for each language was set in order to check whether there was any improvement after more exposure to the streams. After listening to each stream participants performed a two alternative forced choice test where 8 words and 8 part-words were presented. Part-words were also tri-syllabic and

they were formed by taking the last syllable of a word and the first two syllables of another word (*see Table 1*). Participants had to press a button to indicate which of the two alternatives was a recently presented artificial word.

Table 1. Stimuli used in the two word learning tests

	Words	Part-words
Two min language	jasoku	kubifa
	bifale	leRonu
	Ronupi	pipulo
	pulose	sejaso
Four min language	kesujo	jonita
	nitamo	moduga
	dugate	tedife
	difena	nakesu

Four words and part-words were used in the two and four minutes (min) tests. The uppercase indicates that the r is voiced.

4.4. MRI data acquisition

DTI was acquired for each subject in a GE 1.5T scanner. The pulse sequence was single-shot, diffusion-weighted, echo planar acquisition (TR=10000ms; TE=minimum; NEX=1; matrix=256mm x 256mm; FOV=240mm x 240mm; slice thickness=2.5mm; interslice gap=0; in-plane resolution=1mm²; b value= 1000 s/mm²; diffusion gradient directions=50 plus one B0 image). One pair of twins was excluded due to technical issues with the scanner during DWI acquisition for one of the siblings.

4.5. Image processing

The diffusion-weighted images were processed using TBSS (Tract-Based Spatial Statistics, (76)), part of FSL (77). First, FA images were created by fitting a tensor model to the raw diffusion data using FDT, and then brain-extracted using BET (78). All subjects' FA data were then aligned into a common space using the nonlinear registration tool FNIRT (79, 80), which uses a b-spline representation of the registration warp field (81). Next, the mean FA image was created and thinned to create a mean FA skeleton, which represents the centres of all tracts common to the group. Each subject's aligned FA data was then projected onto this skeleton. The latter step was repeated for the remaining diffusivity metrics (MD, AD, RD). Finally, ROIs for the left and right AF, IFOF, and UF were defined based on the JHU white-matter probabilistic atlas (82–84) (*see Figure 3*). The intersection between these ROIs and the FA skeleton was computed, representing the center of the tracts of interest, and the average for each diffusivity measure within the intersected skeleton was computed for each subject.

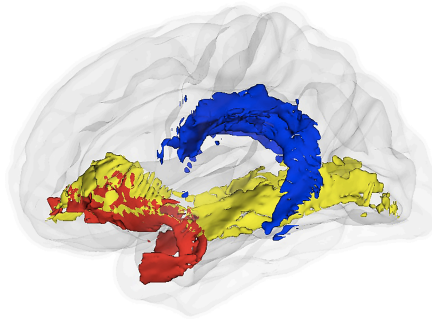


Figure 3. Left hemisphere representation of the intersections between tracts' ROIs from the JHU tractography atlas and the white matter skeleton estimated with the TBSS method. Blue = Arcuate Fasciculus. Yellow = Inferior Fronto-Occipital Fasciculus. Red = Uncinate Fasciculus.

4.6. Statistical analysis

Firstly, the accuracy was calculated for each language. A one-tailed t test against 50 level was performed in order to evaluate whether twins scored above the chance level. Secondly, for each pair of twins, the difference in absolute value divided by the sum was measured for both learning scores (the 2 minute and the 4 minute language). The same procedure was followed for the mean value of the FA, AD, RD and MD for each tractor (AF, IFOF, UF). Finally both measures (behavioral and structural twins' differences) were correlated.

In addition, we also tested whether individual differences in WM were related to their statistical learning ability by computing correlations between WM integrity indices and word learning performance measures.

References

1. Yang J, Gates KM, Molenaar P, Li P (2015) Neural changes underlying successful second language word learning: An fMRI study. *J Neurolinguistics* 33:29-49.
2. López-Barroso D et al. (2013) Word learning is mediated by the left arcuate fasciculus. *Proc Natl Acad Sci U S A* 110:13168–73.
3. Sheppard JP, Wang J-P, Wong PCM (2012) Large-scale cortical network properties predict future sound-to-word learning success. *J Cogn Neurosci* 24:1087–103.
4. McNealy K, Mazziotta JC, Dapretto M (2006) Cracking the language code: neural mechanisms underlying speech parsing. *J Neurosci* 26:7629–39.
5. Cunillera T et al. (2009) Time course and functional neuroanatomy of speech segmentation in adults. *Neuroimage* 48:541–53.
6. Hoff E (2006) How social contexts support and shape language development. *Dev Rev* 26:55–88.
7. Kuhl PK, Tsao F-M, Liu H-M (2003) Foreign-language experience in infancy: effects of short-term exposure and social interaction on phonetic learning. *Proc Natl Acad Sci U S A* 100:9096–101.
8. Johnson J, Newport E (1989) Critical period effects in second language learning: The influence of maturational state on the acquisition of English as a second language. *Cogn Psychol* 21:60–99.

9. Snow C, Hoefnagel-Höhle M (1978) The critical period for language acquisition: Evidence from second language learning. *Child Dev*:1114–1128.
10. Bialystok E, Hakuta K (1999) *Confounded age: Linguistic and cognitive factors in age differences for second language acquisition* ed E-Library T& F (Lawrence Erlbaum Associates, Inc.).
11. Mayberry R, Lock E (2003) Age constraints on first versus second language acquisition: Evidence for linguistic plasticity and epigenesis. *Brain Lang* 87:161–181.
12. DeKeyser R (2000) The robustness of critical period effects in second language acquisition. *Stud Second Lang Acquis* 22:499–533.
13. Linck J, Kroll J, Sunderman G (2009) Losing access to the native language while immersed in a second language: Evidence for the role of inhibition in second-language learning. *Psychol Sci* 20:1507–15.
14. Perani D et al. (2003) The role of age of acquisition and language usage in early, high-proficient bilinguals: an fMRI study during verbal fluency. *Hum Brain Mapp* 19:170–82.
15. Dale PS, Harlaar N, Hayiou-Thomas ME, Plomin R (2010) The etiology of diverse receptive language skills at 12 years. *J Speech Lang Hear Res* 53:982–92.
16. Hayiou-Thomas ME, Dale PS, Plomin R (2012) The etiology of variation in language skills changes with development: a longitudinal twin study of language from 2 to 12 years. *Dev Sci* 15:233–49.

17. Kovas Y et al. (2005) Genetic influences in different aspects of language development: the etiology of language skills in 4.5-year-old twins. *Child Dev* 76:632–51.
18. Van Soelen ILC et al. (2009) Individual differences in dynamic measures of verbal learning abilities in young twin pairs and their older siblings. *Learn Individ Differ* 19:440–444.
19. Chiang M, McMahon K, Zubicaray G de (2011) Genetics of white matter development: a DTI study of 705 twins and their siblings aged 12 to 29. *Neuroimage* 54:2308–2317.
20. Plomin R, DeFries JC, Knopik VS, Neiderhiser JM (2013) *Behavioral Genetics* (Worth Publishers). 6th Ed.
21. Pike A, Reiss D, Hetherington EM, Plomin R (1996) Using MZ differences in the search for nonshared environmental effects. *J Child Psychol Psychiatry* 37:695–704.
22. Turkheimer E, Waldron M (2000) Nonshared environment: a theoretical, methodological, and quantitative review. *Psychol Bull* 126:78–108.
23. Yamagata S et al. (2013) Bidirectional influences between maternal parenting and children's peer problems: a longitudinal monozygotic twin difference study. *Dev Sci* 16:249–59.
24. Caspi A et al. (2004) Maternal expressed emotion predicts children's antisocial behavior problems: using monozygotic-twin differences to identify environmental effects on behavioral development. *Dev Psychol* 40:149–61.
25. Asbury K, Dunn JF, Pike A, Plomin R (2003) Nonshared environmental influences on individual differences in early

behavioral development: a monozygotic twin differences study. *Child Dev* 74:933–43.

26. Deater-Deckard K et al. (2001) Fast-Track Report: Nonshared environmental processes in social-emotional development: an observational study of identical twin differences in the preschool period. *Dev Sci* 4:F1–F6.
27. Barclay NL, Eley TC, Buysse DJ, Maughan B, Gregory AM (2012) Nonshared environmental influences on sleep quality: a study of monozygotic twin differences. *Behav Genet* 42:234–44.
28. Barclay NL, Eley TC, Parsons MJ, Willis TA, Gregory AM (2013) Monozygotic twin differences in non-shared environmental factors associated with chronotype. *J Biol Rhythms* 28:78–108.
29. Liang H, Eley TC (2005) A monozygotic twin differences study of nonshared environmental influence on adolescent depressive symptoms. *Child Dev* 76:1247–60.
30. De Geus EJC et al. (2007) Intrapair differences in hippocampal volume in monozygotic twins discordant for the risk for anxiety and depression. *Biol Psychiatry* 61:1062–71.
31. Van 't Ent D et al. (2007) A structural MRI study in monozygotic twins concordant or discordant for attention/hyperactivity problems: evidence for genetic and environmental heterogeneity in the developing brain. *Neuroimage* 35:1004–20.
32. Den Braber A et al. (2011) White matter differences in monozygotic twins discordant or concordant for obsessive-compulsive symptoms: a combined diffusion tensor

imaging/voxel-based morphometry study. *Biol Psychiatry* 70:969–77.

33. Den Braber A et al. (2012) Brain activation during response interference in twins discordant or concordant for obsessive compulsive symptoms. *Twin Res Hum Genet* 15:372–83.
34. Duffau H (2006) Brain plasticity: from pathophysiological mechanisms to therapeutic applications. *J Clin Neurosci* 13:885–897.
35. Duffau H (2008) The anatomo-functional connectivity of language revisited. New insights provided by electrostimulation and tractography. *Neuropsychologia* 46:927–34.
36. Catani M et al. (2013) A novel frontal pathway underlies verbal fluency in primary progressive aphasia. *Brain* 136:2619–28.
37. Chevillet M, Jiang X (2013) Automatic phoneme category selectivity in the dorsal auditory stream. *J Neurosci* 33:5208–5215.
38. Rolheiser T (2011) Dynamic processing in the human language system: synergy between the arcuate fascicle and extreme capsule. *J Neurosci* 31:1193–1212.
39. Buchsbaum B, D’Esposito M (2008) The search for the phonological store: from loop to convolution. *J Cogn Neurosci* 20:762–778.
40. Schulze K (2012) Test of a motor theory of long-term auditory memory. *Proc Natl Acad Sci U S A* 109:7121–7125.

41. Romanski L, Tian B, Fritz J (1999) Dual streams of auditory afferents target multiple domains in the primate prefrontal cortex. *Nat Neurosci* 2:1131–1136.
42. Mestres-Missé A (2008) Functional neuroanatomy of meaning acquisition from context. *J Cogn Neurosci* 20:2153–2166.
43. Hickok G, Poeppel D (2007) The cortical organization of speech processing. *Nat Rev Neurosci* 8:393–402.
44. Epelbaum S, Pinel P, Gaillard R, Delmaire C (2008) Pure alexia as a disconnection syndrome: new diffusion imaging evidence for an old concept. *Cortex* 44:962–974.
45. Catani M, Mesulam M (2008) The arcuate fasciculus and the disconnection theme in language and aphasia: history and current state. *Cortex* 44:953–61.
46. McDonald C, Ahmadi M, Hagler D (2008) Diffusion tensor imaging correlates of memory and language impairments in temporal lobe epilepsy. *Neurology* 71:1869–1876.
47. McDonald CR et al. (2014) White matter microstructure complements morphometry for predicting verbal memory in epilepsy. *Cortex* 58:139–150.
48. Walsh M et al. (2011) Object working memory performance depends on microstructure of the frontal-occipital fasciculus. *Brain Connect* 1:317–29.
49. Sepulcre J, Masdeu J, Sastre-Garriga J (2008) Mapping the brain pathways of declarative verbal memory: evidence from white matter lesions in the living human brain. *Neuroimage* 42:1237–1243.

50. Petrides M, Alivisatos B, Meyer E, Evans AC (1993) Functional activation of the human frontal cortex during the performance of verbal working memory tasks. *Proc Natl Acad Sci* 90:878–882.
51. Cappell KA, Gmeindl L, Reuter-Lorenz PA (2010) Age differences in prefrontal recruitment during verbal working memory maintenance depend on memory load. *Cortex* 46:462–73.
52. Saffran JR, Aslin RN, Newport EL (1996) Statistical learning by 8-month-old infants. *Science* 274:1926–1928.
53. Westlye L, Walhovd K, Dale A (2009) Life-span changes of the human brain white matter: diffusion tensor imaging (DTI) and volumetry. *Cereb Cortex* 20:2055–68.
54. Tamnes C, Østby Y, Fjell A (2010) Brain maturation in adolescence and young adulthood: regional age-related changes in cortical thickness and white matter volume and microstructure. *Cereb Cortex* 20:534–548.
55. Gao W, Lin W, Chen Y (2009) Temporal and spatial development of axonal maturation and myelination of white matter in the developing brain. *Am J Neuroradiol* 30:290–296.
56. Ishibashi T et al. (2006) Astrocytes promote myelination in response to electrical impulses. *Neuron* 49:823–32.
57. Markham J, Greenough W (2004) Experience-driven brain plasticity: beyond the synapse. *Neuron Glia Biol* 1:351–363.
58. Draganski B, May A (2008) Training-induced structural changes in the adult human brain. *Behav Brain Res* 192:137–142.

59. Lessov-Schlaggar CN et al. (2012) Longitudinal genetic analysis of brain volumes in normal elderly male twins. *Neurobiol Aging* 33:636–44.
60. Geng X et al. (2012) White matter heritability using diffusion tensor imaging in neonatal brains. *Twin Res Hum Genet* 15:336–50.
61. Kochunov P, Jahanshad N, Sprooten E (2014) Multi-site study of additive genetic effects on fractional anisotropy of cerebral white matter: Comparing meta and megaanalytical approaches for data pooling. *Neuroimage* 95:136–150.
62. Beaulieu C (2002) The basis of anisotropic water diffusion in the nervous system - a technical review. *NMR Biomed* 15:435–55.
63. Alexander AL, Lee JE, Lazar M, Field AS (2007) Diffusion tensor imaging of the brain. *Neurotherapeutics* 4:316–29.
64. Hauser MD, Newport EL, Aslin RN (2001) Segmentation of the speech stream in a non-human primate: statistical learning in cotton-top tamarins. *Cognition* 78:B53–64.
65. Toro JM, Trobalon JB, Sebastián-Gallés N (2005) Effects of Backward Speech and Speaker Variability in Language Discrimination by Rats. *J Exp Psychol Anim Behav Process* 31:95–100.
66. Thiebaut de Schotten M, Dell’Acqua F, Valabregue R, Catani M (2012) Monkey to human comparative anatomy of the frontal lobe association tracts. *Cortex* 48:82–96.
67. Brauer J, Anwander A, Friederici A (2011) Neuroanatomical prerequisites for language functions in the maturing brain. *Cereb Cortex* 21:459–466.

68. Poldrack R, Sabb F (2005) The neural correlates of motor skill automaticity. *J Neurosci* 25:5356–5364.
69. Doyon J, Penhune V, Ungerleider L (2003) Distinct contribution of the cortico-striatal and cortico-cerebellar systems to motor skill learning. *Neuropsychologia* 41:252–262.
70. Saint-Cyr J (2003) Frontal-striatal circuit functions: context, sequence, and consequence. *J Int Neuropsychol Soc* 9:103–128.
71. Graaf F Van der, Jong B De (2004) Cerebral activation related to skills practice in a double serial reaction time task: striatal involvement in random-order sequence learning. *Cogn Brain Res* 20:120–131.
72. Dutoit I, Pagel C, Pierret N, Baraille E (1996) The MBROLA project = Towards a set of high quality speech synthesizers free of use for non commercial purposes. 3:1393–1396.
73. Aslin R, Saffran J, Newport E (1998) Computation of conditional probability statistics by 8-month-old infants. *Psychol Sci* 9:321–324.
74. Johnson E, Jusczyk P (2001) Word segmentation by 8-month-olds: When speech cues count more than statistics. *J Mem Lang* 44:548–567.
75. Saffran J, Newport E, Aslin R (1996) Word segmentation: The role of distributional cues. *J Mem Lang* 35:606–621.
76. Smith SM et al. (2006) Tract-based spatial statistics: voxelwise analysis of multi-subject diffusion data. *Neuroimage* 31:1487–505.

77. Smith SM et al. (2004) Advances in functional and structural MR image analysis and implementation as FSL. *Neuroimage* 23 Suppl 1:S208–19.
78. Smith SM (2002) Fast robust automated brain extraction. *Hum Brain Mapp* 17:143–55.
79. Andersson JL, Jenkinson M, Smith S (2007) Non-linear optimisation. FMRIB Technical Report TR07JA1.
80. Andersson JL, Jenkinson M, Smith S (2007) Non-linear registration, aka Spatial normalisation. FMRIB technical report TR07JA2.
81. Rueckert D et al. (1999) Nonrigid registration using free-form deformations: application to breast MR images. *IEEE Trans Med Imaging* 18:712–21.
82. S. Mori, S. Wakana, L.M. Nagae-Poetscher and PCM van Z (2005) *MRI Atlas of Human White Matter* ed Elsevier (Amsterdam).
83. Wakana S et al. (2007) Reproducibility of quantitative tractography methods applied to cerebral white matter. *Neuroimage* 36:630–44.
84. Hua K et al. (2008) Tract probability maps in stereotaxic spaces: analyses of white matter anatomy and tract-specific quantification. *Neuroimage* 39:336–47.

4. GENERAL DISCUSSION

The main goal of this dissertation is to investigate to what extent genetic and environmental factors explain the neural underpinnings of word and rule learning. To accomplish this goal we performed two studies. In the first study, we analyzed the influence of genetic sources on brain activity elicited after exposure to an artificial language. The idea behind this study was to know whether or not MZT were more similar than DZT when brain signal was compared between them. The goal of the second study was to understand how experience shapes WM structures supporting word learning. In this case we analyzed if differences between MZT in language-related WM structures were associated with differences in word segmentation ability. These two objectives contribute to the ongoing debate around the origin of individual differences on language learning. The present dissertation intends to contribute to a better understanding on the extent to which the different structures and mechanism(s) allowing both word and rule learning are inherited and acquired.

Here, I will first summarize the findings and I will extend the discussion of the two studies presented. Second, I will expose the implications of the findings with respect to the literature on language learning. Then, I will explain how the methodology applied can help the improvement of subsequent twin studies, and finally, I will conclude by proposing future lines of research.

4.1. Summary of results and discussion

4.1.1. Heritability in both word and rule learning

In order to investigate the influence of genetics and environment on both word and rule learning, ERPs were measured in a sample of MZT and DZT while they were listening to four artificial languages containing AXC rules (eg., **pesedi**, **pemudi**, **pefodi**). The procedure we followed was previously applied in De Diego Balaguer et al., (2007), Peña et al. (2002) among others.

As a first step, we performed an ERP analysis in order to confirm and extend the learning effects previously reported using similar paradigms. The results from this first analysis showed an increase in amplitude of the N4 and P2 from the first to the fourth minute of exposure to the languages. These findings are in accordance with previous research in which N4 and P2 have been associated with the learning of words and rules, respectively (Cunillera et al., 2009, 2006; Sanders et al., 2002). By taking this step, we also made sure that any differences between MZT and DZT (when their sibling-ship and zygozity were controlled for) could explain subsequent results.

Afterwards, a wavelet analysis was applied and brain activity was compared between twin pairs for the two time windows corresponding to the appearance of the two neural signatures (N4:

350-450 ms and P2: 150-250 ms) and in the first and fourth minute of exposure to the languages. Furthermore, siblings' similarity was compared between groups, and heritability was measured using Falconer's formula (Falconer & Mackay, 1996) for both word and rule learning in the fourth minute of exposure.

Results showed that MZT were more similar compared to DZT in both windows and both minutes. However, heritability indicated a low effect of genetic factors in both types of learning (*see figure 8*). These findings suggest that genetics plays a role when explaining individual differences on word and rule learning, although this role is not of significance when compared with the effects of environment. As hypothesized, heritability for word learning was higher than for rule learning (i.e., word learning: 0.189; rule learning: 0.163), but compared with previous studies in which children were tested (Dale et al., 2000; Dionne et al., 2003; Hayiou-Thomas et al., 2012; Kovas et al., 2005), our indices were much lower.

We cannot discard the possibility that our low heritability values could be driven by the lack of statistical power, since the sample size we used was smaller when compared with other ERP twin studies (Visscher, 2004). However, since differences between groups were observed, this assumption would be improbable.

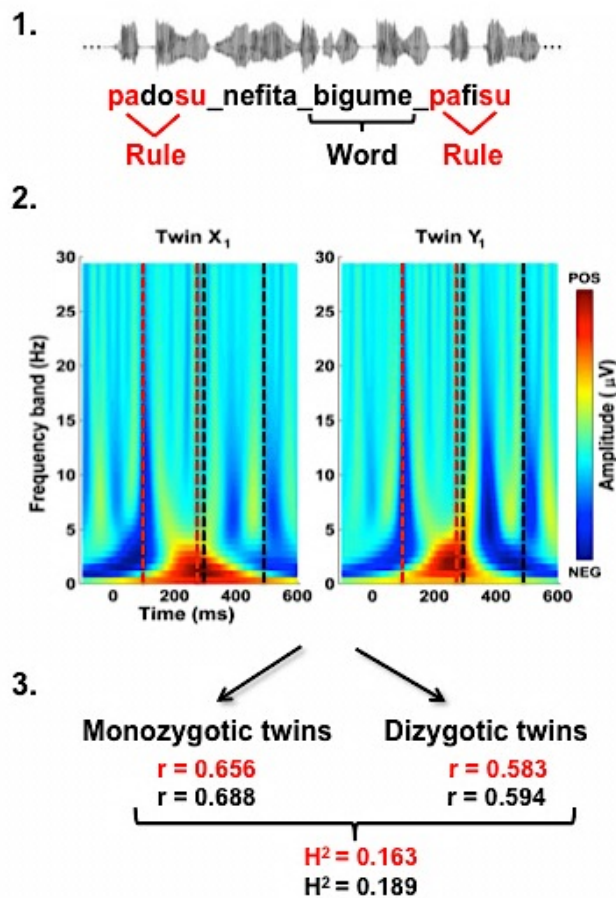


Figure 8. Summary of the experimental procedure and results for the first study. (1) Twins listened to a sequence of words following an AXC rule. (2) From the EEG obtain during the perception of the speech stream a wavelet analysis was applied. (3) The cross-correlation (r) for each pair of twins was measured in the time window associated with the processing of rules (red) and words (black). Afterwards, the heritability was calculated (H^2).

Previous studies with infants and children reported a moderate effect of genetics for both vocabulary and grammar (between 0.3 and 0.5) (Byrne et al., 2006; Yulia Kovas et al., 2005). Importantly, significant differences can be established between other language twin studies and ours. First of all, they tested young twins who were

evaluated through reports of their parents or standardized tests, which required an explicit answer from them (e.g., Hayiou-Thomas et al., 2012). Therefore, in most of these studies, children, had to first understand the question or task, and second, they had to verbalize or write the answer. In order to do so, it is not only the language areas (for both perception and production) that were involved, but other non-language areas as well, such as those related to memory, attention, motor and visual processing (in those tasks where they had to write, read or see and analyse an object). This is extremely important when studying heritability because it makes the measure very imprecise. Therefore, one possible explanation for the difference between our heritability values and the ones observed in other language-related twin studies is that our study is based on simpler processes significantly constrained to the task of learning both words and rules.

Another possible explanation for our low values might be related to the age of our participants. That is, heritability seems to decrease as function of age, due to the increase of both language and non-language experiences (Zendel & Alain, 2012). Indeed, Batouli, Trollor, Wen, & Sachdev (2014) reviewed the heritability of brain volumes from birth to early adulthood and they observed an increment of heritability until the age of 20 and a decrement later in life. Since most of our participants were older than 20, the results of Batouli et al. (2014) would be in accordance with ours. Importantly, studies testing vocabulary and grammar in different age groups (e.g., from 2 to 12), reported an increase of heritability for both

abilities across ages (Hayiou-Thomas et al., 2012; Hayiou-Thomas, Harlaar, Dale, & Plomin, 2010). Hence, our findings would complement these twin studies by extending the age range of participants and showing a decrease of heritability in early adulthood.

The small difference in heritability found between word and rule learning could be explained by the late maturation of frontal areas associated with attention (Sowell et al., 2004). Importantly, the effect observed in the P2 window has been associated with a reallocation of attention towards the extraction of rules (De Diego Balaguer et al., 2007). Hence, the slower development of frontal areas might make the attentional system more susceptible to the effect of environmental factors (Posner, Sheese, Odludaş, & Tang, 2006).

Summing up, the higher similarity observed in MZT indicates that, as we expected, the influence of genetics can be observed during word and rule learning, although its magnitude (in terms of heritability) is low for both types of processing, but higher for the word learning one.

4.1.2. Environmental contribution to brain structures associated with statistical language learning

In order to study how environment affects WM structures involved in statistical learning we selected a sample of MZT who were then tested on their statistical learning ability. Specifically, they were exposed to two artificial language streams before a test. One stream was two minutes long, and the other was four minutes long, both formed by four nonsense words. Participants also underwent MRI acquisition. Diffusion tensor images were acquired and processed, and several diffusivity metrics (FA, RD, AD, MD) were extracted for each of the WM tracts previously associated with language processing (AF, IFOF, UF) (Catani et al., 2002, 2005; López-Barroso et al., 2013).

Our analysis was based on the following assumption: since MZT share (almost) 100% of their DNA, as well as their pre- and post-natal environment, any difference between them would most likely be caused by non-shared environmental factors. Thus, we calculated the difference between twins in both the behavioural and the WM measures and then we correlated these two measures to see if dissimilarities in WM were associated with dissimilarities in word learning.

Similarity in the integrity of two of the three tracts (left AF and left IFOF) correlated with similarity in accuracy values in the 2- and 4-minutes languages streams. Specifically, in the left AF, both FA and

AD similarity correlated with similarity in performance after 2 minutes of language exposure, whereas in left IFOF, similarity in all four diffusivity measures (FA, AD, RD and MD) correlated with performance similarity after exposure to the 4-minute language stream (*see figure 9*).

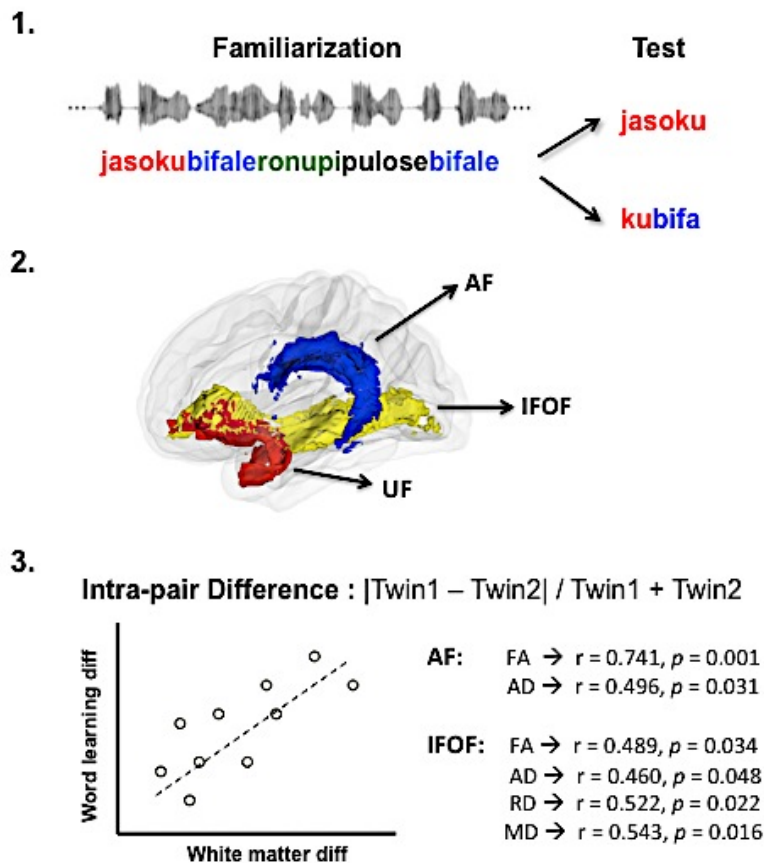


Figure 9. Summary of the experimental procedure and results for the second study. (1) Twins listened to a speech stream and they performed a test on word learning. (2) White matter integrity measures (fractional anisotropy (FA), axial diffusivity (AF), radial diffusivity (RD) and mean diffusivity (MD)) for the arcuate fasciculus (AF), the inferior fronto-occipital fasciculus (IFOF) and the uncinate fasciculus (UF) were obtained for each twin. (3) Intra-pair differences in the word learning test and the white matter structures were correlated. Significant results were found in the AF and IFOF.

Regarding the findings in the AF, it should be noted that correlations were observed only in the short language stream. The AF connects language receptive areas with the premotor/motor areas, reaching Broca's area. Together with the superior longitudinal fasciculus, the AF is one of the main routes in the language dorsal pathway (Axer, Klingner, & Prescher, 2013). This pathway has an auditory-motor integration function, which is essential in decoding the signal into the phonetic structure of words and transforming it into articulatory representations (Hickok & Poeppel, 2007). As mentioned in the introduction (p. 19), fMRI studies on statistical learning found higher activation in the premotor area when words were compared with part-words/non-words (Cunillera et al., 2009) or when the perception of a speech stream was compared with a random syllable stream (McNealy et al., 2006). Hence, our results would be in agreement with previous functional studies on word segmentation by showing that environmental changes in those connections between the auditory and motor areas would be, to a certain extent, responsible for individual differences observed in word segmentation.

As for the IFOF, our findings indicate a relationship between twin differences in this tract for all structural measures and twin differences in performance after familiarization with the 4-minutes speech stream. The IFOF is a tract extending from the occipital to the frontal areas, passing through the temporal lobe. It has been considered as part of the language ventral pathway (Dick & Tremblay, 2012), and it is characterized by its function in working

memory (McDonald, Ahmadi, & Hagler, 2008; McDonald et al., 2014; Walsh et al., 2011). A recent case study investigated statistical learning in a patient with middle temporal damage. The results showed a significant deficit in his ability to extract statistical regularities (Schapiro et al., 2014). Therefore, it is possible that IFOF, through its connections between the prefrontal cortex – crucial for verbal working memory (Cappell, Gmeindl, & Reuter-Lorenz, 2010; Petrides, Alivisatos, Meyer, & Evans, 1993) – and the inferior temporal cortex, would support the maintenance of the new information in memory. Indeed, behavioural results would be aligned with this argument since higher accuracy was found in the 4-minutes language test when compared with the 2-minutes one. In addition, these results were found only in the left hemisphere, suggesting that WM experience differences were mainly language-related.

In summary, results from this second study point in the direction that individual differences observed in statistical learning are related to environmental changes in both the AF and the IFOF.

4.2. Implications on the language learning literature

4.2.1. Individual differences in second language learning

Individual differences in adult language processing so far have been mainly studied in L2 learning (Bates, Dale, & Tal, 1995; Dörnyei, 2005; Ellis, 2004; Farmer, Misyak, & Christiansen, 2012; MacDonald & Christiansen, 2002; Michael & Gollan, 2005; Vasilyeva, Waterfall, & Huttenlocher, 2008). However, some studies have also explored individual differences in the underlying mechanisms allowing the acquisition of language. More specifically, some studies focusing on word and rule learning have reported individual differences in behaviour (Arciuli & Simpson, 2011; Saffran, 2001) as well as in brain activity and structure (Cunillera et al., 2009, 2006; De Diego Balaguer et al., 2007; McNealy et al., 2006; Sanders et al., 2002).

As presented in the introduction of this dissertation (p. 25), one of the main factors determining individual differences in L2 learning is age of acquisition. One relevant model that has explored how age of acquisition affects vocabulary and grammar learning is Ullman's model of declarative/procedural memory (Ullman, 2001, 2005). According to this model, language learning depends on two memory systems. One is *declarative memory* that has been related to the learning of arbitrary units of information in an explicit manner (Eichenbaum, 2000). This system would be mainly

supported by the hippocampus and medial temporal lobe and would be related to the acquisition of vocabulary, an ability that does not seem to be dependent on age. The other system is the *procedural memory* that has been associated with implicit and sequential learning. It would be processed through networks involving frontal areas and basal ganglia and would be associated with rule learning. This type of learning would be more restricted by age when the acquisition of a new language is taking place.

At first sight, Ullman's model seems to resonate with the approach of our studies; however, certain aspects need to be specified. When defining the model, Ullman associated vocabulary learning with the acquisition of isolated bits of information (i.e., words). However, in our approach words were not isolated but concatenated (with and without subliminal pauses of 25 ms). Indeed, word learning in our approach implied the extraction of sequential information (i.e., transitional probabilities), which would be processed implicitly. Moreover, some studies have found activity in the basal ganglia during tasks with similar characteristics as ours (Karuza et al., 2013). Therefore, it seems that procedural memory would underlie not only rule learning, but also word learning (at least in the first learning stages). This idea is in agreement with our results showing that genetic pressures (as well as environmental ones) similarly affect both word and rule learning. Our findings would therefore suggest that both learning processes would be affected by age in similar way. So far, there is a lack of models describing how mechanisms of word and rule learning of the type analysed here

impact on L2 learning. This is a line of research to be developed in the future.

Finally, our second study also suggests that subsequent learning models must take WM into consideration in addition to grey matter and subcortical structures when explaining individual differences in L2 learning and when defining age restrictions of the acquisition of the different language aspects.

4.2.2. Word and rule extraction: one or two mechanisms?

One unsolved debate in the language learning and processing literature revolves around whether there is a single domain-general mechanism underlying language learning or if more than one mechanism is required. Although this was not at the core of the present dissertation, we believe some comments are in due course here. According to the first hypothesis, associative mechanisms, such as statistical learning, would suffice to explain both word and rule learning (i.e. Altmann, 2002; Elman, 2005; Elman, 2006; Perruchet et al., 2006, 2004; Seidenberg, 1997). In the second case, researchers argue the need for two different mechanisms: one for word learning and another for grammar (rule) learning (Bonatti et al., 2006; Marcus, 1999; Peña et al., 2002). For example, the *More than One Mechanism Hypothesis* (Endress & Bonatti, 2007) postulates that, even though learners can compute statistical relations, they do not use this ability to extract rules.

In our first study, when word and rule learning were compared, no differences were observed in terms of the estimated magnitude of genetic/environmental influences. This is, to some extent, consistent with Dale et al. (2000) and Dionne et al. (2003) in which high genetic correlations were found between vocabulary and grammar. These authors suggested that the same genetic sources would underlie both types of learning.

Our findings can be interpreted in three ways. First, word and rule learning may have the same genetic and/or environmental origin (Dale et al., 2000; Dionne et al., 2003), which would reflect the use of the same neural substrates and, therefore, supporting a single underlying mechanism. Going against this account, however, previous studies have observed different brain structures involved in these two processes (Cunillera et al., 2009; Forkstam et al., 2006; McNealy et al., 2006; Petersson, 2004) which, at the same time, would be supported by different genetic sources (Bearden et al., 2012; Wong, Morgan-Short, Ettlinger, & Zheng, 2012). A second possibility is that both processes are sustained by different mechanisms in which the pressures of genetics and/or environment would be similar. Finally, and as previously mentioned, a third explanation for the absence of genetic/environmental differences between conditions could be driven by the lack of statistical power. Even though the analysis we used was able to identify differences between twin types, it might not be strong enough to uncover potential differences between processes. Hence, maybe the small

effects observed could be enhanced with an increase of the sample size. All three hypotheses should be further investigated in order to reach a clearer conclusion.

4.3. New methodological approaches

One of the main innovations of this dissertation is the implementation of original methods. Regarding our first study, one of our main methodological challenges was to find the optimal way to compare twins' electrophysiological measurements. Our aim was to evaluate how similar the brain activity from two individuals was. Our first approach was to select the mean amplitude from the components we were interested in (P2 and N4) for each participant and to measure the difference between them. However, the high time resolution of the signal created several problems. For instance, in some cases the pattern of the signal was highly similar between a pair of twins but was shifted some milliseconds for one sibling with respect to the other, resulting in misleading similarity estimates. By transforming the signal into a wavelet map, there were some advantages. First, we could take a third characteristic of the signal into account, namely, its frequency. Second, by measuring the maximum cross-correlation, we were able to correct for possible shifts existing between twin pairs, thus ensuring that the highest value of intra-pair similarity would be considered for further analyses. Though noise was still present in the data, the wavelet cross-correlations allowed us to observe differences across groups,

thus showing that the method used could be effective when analysing brain activity in small samples of twins.

Note that the goal of our second study was to evaluate the effect of environment on the brain structures associated with word learning. As explained in the introduction of this dissertation (p. 29), two methods have been used in previous research to accomplish goals similar to ours (i.e., studying the environmental effects on specific phenotypes). One of these methods is the discordant twin design consisting of the selection of MZT pairs with maximally dissimilar scores. These twins are subsequently compared in their brain structure (de Geus et al., 2007; den Braber et al., 2011, 2012; van 't Ent et al., 2007). In the second method, called the MZT difference design, instead of creating groups of highly discordant twins, discordancy is allowed to vary in a continuous (non discrete) manner, so that a spectrum of intra-pair variation (ranging from very similar to very dissimilar) can be included in the sample (Pike et al., 1996; Plomin et al., 2013; Turkheimer & Waldron, 2000). Here, we used the rationale of the discordant twins design and combined it with the kind of measures used in the MZT difference design. In other words, we computed the differences between MZT siblings in the WM structures of interest and correlated them with the differences observed in word learning performance. Importantly, the advantage here is that neither a large sample nor a special one (i.e., extremely discordant MZT) is needed. To our knowledge, no previous research has applied the MZT difference strategy to study either brain structure or function. Hence, this approach could help

researchers interested in using neuroimaging to investigate the effect of experience in the brain.

In short, this dissertation not only provided new information about the origin of individual differences in language learning but also offered a new perspective on how twins' data can be analysed by using both electrophysiological and neuroimaging techniques.

4.4. Future lines of research

Although our results shed some light on the importance of environmental components that mould basic language learning mechanisms, it is important to note that adult twins formed our sample, hence we cannot make inferences about how word and rule learning are shaped by nature and nurture components in infants who are still acquiring their L1. Thus, an open question for further research would be to investigate how both genetic and environmental factors influence word and rule learning in infancy. Would genetic sources have a stronger effect? Or would environment already be the principle factor responsible for language learning variance? One way to address these questions would be to evaluate twin infants with simplified versions of the speech sequences used by De Diego Balaguer et al. (2007) (like, for instance, Gómez & Maye, 2005) and to measure the similarity between siblings across time at both behavioural and neural levels. Future research may also explore other issues raised by the present work. Focusing on our first study, it is important to note that the

rules studied were more like to morphological rules (e.g., **unbelievable**, **untreatable**) as opposed to syntactic rules. That is, our stimuli were formed by AXC rules where each element was a syllable, and the ERP signatures we observed were more related to morphological as opposed to syntactic processing (Domínguez, Alija, Cuetos, & de Vega, 2006; Domínguez, de Vega, & Barber, 2004). The use of other types of stimuli might result in syntax-type processing. For instance, Mueller et al. (2008) used dissyllabic stimuli also formed by structures following AXC rules and observed syntax-related ERP responses (i.e., P600 and LAN). Therefore, it would be interesting to investigate how brain activity elicited by this type of rule would be affected by both genetic and environmental influences. Considering that the processes for both morphologic and syntactic rules are based on the extraction of similar structures, we should expect to observe the same influence of genetics and environment in both cases.

There are other aspects of L2 learning in which individual differences have been also observed. One clear example is phoneme perception (Díaz, Baus, Escera, Costa, & Sebastián-Gallés, 2008; Jakoby, Goldstein, & Faust, 2011; Jin, Díaz, Colomer, & Sebastián-Gallés, 2014; Sebastián-Gallés et al., 2012). Some research has shown that many L2 learners, despite having an early exposure to an L2 (2-4 years old) and being extensively exposed to it, still experience serious difficulties when perceiving certain L2 speech sounds, while others can behave as native speakers (Sebastián-Gallés & Baus, 2005). These differences across individuals have

been attributed to language-specific mechanisms rather than general-acoustic ones (Díaz et al., 2008), and they seem to be related to L1 perceptual abilities (Díaz et al., 2008; Díaz, Sebastián-Gallés, Erdocia, Mueller, & Laka, 2011). The existence of a link between L1 and L2 phoneme processing makes the investigation of the role of genetic and environmental factors in such individual differences an attractive topic. In the annex of this dissertation a summary of a study following this idea is presented.

Finally, in our second study only WM integrity was analysed. A sensible subsequent step in this project is the study of how environmental modifications in grey matter (at the cortex and at basal ganglia) can explain individual differences in statistical learning. Although there is no previous research in which the structure of grey matter and/or subcortical structures have been associated with statistical learning, some hypotheses can be formulated based on previous functional studies. At the cortical level, the left premotor area together with the left inferior parietal cortex, the superior temporal gyrus and, specially, the premotor cortex would deserve particular attention (Cunillera et al., 2009; López-Barroso et al., 2013; McNealy et al., 2006). At the subcortical level, we propose that the putamen and the caudate nucleus would be the structures that contribute the most to the statistical learning variability (Karuza et al., 2013; McNealy et al., 2006).

5. SUMMARY AND CONCLUSIONS

Based on the results obtained in this dissertation, several conclusions can be reached. From our first study we conclude that:

- Genetic factors are salient when explaining individual differences in word and rule learning ability.
- The low heritability indices found in both word and rule learning indicate that environment would be the principal factor that defines these two processes.
- The lack of differences between word and rule learning in terms of genetic and environmental influences suggest that either both mechanisms are underlying the same factors or both processes are part of the same mechanism.

From our second study we conclude that:

- Experience molds the structure of the left AF and this would likely be associated with the way words are segmented and kept in working memory.
- Experience also shapes the structure of left IFOF in a way that affects how words are consolidated in memory.

Taken together, the results of this dissertation suggest that some of the basic mechanisms required when learning a language are affected by our previous experiences. Moreover, the changes from

these external sources would be reflected in our brain activity and in specific WM structures (AF and IFOF) associated with language processing.

6. REFERENCES

- Abla, D., Katahira, K., & Okanoya, K. (2008). On-line Assessment of Statistical Learning by Event-related Potentials. *Journal of Cognitive Neuroscience*, 20(6), 952–64.
- Alain, C., Campeanu, S., & Tremblay, K. (2010). Changes in sensory evoked responses coincide with rapid improvement in speech identification performance. *Journal of Cognitive Neuroscience*, 22(2), 392–403.
- Altmann, G. T. M. (2002). Learning and development in neural networks – the importance of prior experience. *Cognition*, 85(2), B43–B50.
- Arciuli, J., & Simpson, I. C. (2011). Statistical learning in typically developing children: the role of age and speed of stimulus presentation. *Developmental Science*, 14(3), 464–73.
- Arciuli, J., Torkildsen, J. von K., Stevens, D. J., & Simpson, I. C. (2014). Statistical learning under incidental versus intentional conditions. *Frontiers in Psychology*, 5, 747.
- Asbury, K., Dunn, J. F., Pike, A., & Plomin, R. (2003). Nonshared environmental influences on individual differences in early behavioral development: a monozygotic twin differences study. *Child Development*, 74(3), 933–43.
- Aslin, R. N. R., Saffran, J. R., & Newport, E. L. EL. (1998). Computation of conditional probability statistics by 8-month-old infants. *Psychological Science*, 9(4), 321–324.
- Axer, H., Klingner, C. M., & Prescher, A. (2013). Fiber anatomy of dorsal and ventral language streams. *Brain and Language*, 127(2), 192–204.
- Badzakova-Trajkov, G., Häberling, I. S., & Corballis, M. C. (2010). Cerebral asymmetries in monozygotic twins: an fMRI study. *Neuropsychologia*, 48(10), 3086–93.
- Bahlmann, J., Gunter, T. C., & Friederici, A. D. (2006). Hierarchical and linear sequence processing: an electrophysiological exploration of

- two different grammar types. *Journal of Cognitive Neuroscience*, 18(11), 1829–42.
- Barclay, N. L., Eley, T. C., Buysse, D. J., Maughan, B., & Gregory, A. M. (2012). Nonshared environmental influences on sleep quality: a study of monozygotic twin differences. *Behavior Genetics*, 42(2), 234–44.
- Barclay, N. L., Eley, T. C., Parsons, M. J., Willis, T. A., & Gregory, A. M. (2013). Monozygotic twin differences in non-shared environmental factors associated with chronotype. *Journal of Biological Rhythms*, 28(1), 78–108.
- Bates, E. A., Friederici, A. D., Wulfeck, B. B., & Juarez, L. A. (1988). On the preservation of word order in aphasia: Cross-linguistic evidence. *Brain and Language*, 33(2), 323–364.
- Batouli, S. A. H., Trollor, J. N., Wen, W., & Sachdev, P. S. (2014). The heritability of volumes of brain structures and its relationship to age: a review of twin and family studies. *Ageing Research Reviews*, 13, 1–9.
- Bearden, C. E., Karlsgodt, K. H., Bachman, P., van Erp, T. G. M., Winkler, A. M., & Glahn, D. C. (2012). Genetic architecture of declarative memory: implications for complex illnesses. *The Neuroscientist: A Review Journal Bringing Neurobiology, Neurology and Psychiatry*, 18(5), 516–32.
- Bettegowda, A., & Smith, G. W. (2007). Mechanisms of maternal mRNA regulation: implications for mammalian early embryonic development. *Frontiers in Bioscience: A Journal and Virtual Library*, 12, 3713–26.
- Bialystok, E., & Hakuta, K. (1999). Confounded age: Linguistic and cognitive factors in age differences for second language acquisition. (T. & F. E-Library, Ed.) *Second language acquisition and the ...*. Lawrence Erlbaum Associates, Inc.
- Bishop, D. V. M., Adams, C. V., & Norbury, C. F. (2006). Distinct genetic influences on grammar and phonological short-term memory deficits: evidence from 6-year-old twins. *Genes, Brain, and Behavior*, 5(2), 158–69.

- Bishop, D. V., Bishop, S. J., Bright, P., James, C., Delaney, T., & Tallal, P. (1999). Different origin of auditory and phonological processing problems in children with language impairment: evidence from a twin study. *Journal of Speech, Language, and Hearing Research: JSLHR*, 42(1), 155–68.
- Bishop, D. V., Chan, J., Adams, C., Hartley, J., & Weir, F. (2000). Conversational responsiveness in specific language impairment: evidence of disproportionate pragmatic difficulties in a subset of children. *Development and Psychopathology*, 12(2), 177–99.
- Bonatti, L. L., Peña, M., Nespor, M., & Mehler, J. (2006). How to hit Scylla without avoiding Charybdis: comment on Perruchet, Tyler, Galland, and Peereeman (2004). *Journal of Experimental Psychology. General*, 135(2), 314–21; discussion 322–6.
- Boomsma, D., Busjahn, A., & Peltonen, L. (2002). Classical twin studies and beyond. *Nature Reviews. Genetics*, 3(11), 872–82.
- Boulinier, T., & Staszewski, V. (2008). Maternal transfer of antibodies: raising immuno-ecology issues. *Trends in Ecology & Evolution*, 23(5), 282–8.
- Byrne, B., Coventry, W. L., Olson, R. K., Samuelsson, S., Corley, R., Willcutt, E. G., ... DeFries, J. C. (2009). Genetic and Environmental Influences on Aspects of Literacy and Language in Early Childhood: Continuity and Change from Preschool to Grade 2. *Journal of Neurolinguistics*, 22(3), 219–236.
- Byrne, B., Olson, R. K., Samuelsson, S., Wadsworth, S., Corley, R., DeFries, J. C., & Willcutt, E. (2006). Genetic and environmental influences on early literacy. *Journal of Research in Reading*, 29(1), 33–49.
- Byrne, B., Olson, R. K., Samuelsson, S., Wadsworth, S., Corley, R., DeFries, J. C., & Willcutt, E. (2013). Genetic and environmental influences on early literacy. *Journal of Research in Reading*, 29(1), 33–49.
- Cappell, K. A., Gmeindl, L., & Reuter-Lorenz, P. A. (2010). Age differences in prefrontal recruitment during verbal working memory maintenance depend on memory load. *Cortex*, 46(4), 462–73.

- Cartmill, E. A., Armstrong, B. F., Gleitman, L. R., Goldin-Meadow, S., Medina, T. N., & Trueswell, J. C. (2013). Quality of early parent input predicts child vocabulary 3 years later. *Proceedings of the National Academy of Sciences of the United States of America*, 110(28), 11278–83.
- Caspi, A., Moffitt, T. E., Morgan, J., Rutter, M., Taylor, A., Arseneault, L., ... Polo-Tomas, M. (2004). Maternal expressed emotion predicts children's antisocial behavior problems: using monozygotic-twin differences to identify environmental effects on behavioral development. *Developmental Psychology*, 40(2), 149–61.
- Catani, M., Howard, R. J., Pajevic, S., & Jones, D. K. (2002). Virtual in Vivo Interactive Dissection of White Matter Fasciculi in the Human Brain. *NeuroImage*, 17(1), 77–94.
- Catani, M., Jones, D. K. D., & Ffytche, D. H. (2005). Perisylvian language networks of the human brain. *Annals of Neurology*, 57(1), 8–16.
- Catani, M., & Mesulam, M. (2008). The arcuate fasciculus and the disconnection theme in language and aphasia: history and current state. *Cortex; a Journal Devoted to the Study of the Nervous System and Behavior*, 44(8), 953–61.
- Catani, M., & Thiebaut de Schotten, M. (2008). A diffusion tensor imaging tractography atlas for virtual in vivo dissections. *Cortex; a Journal Devoted to the Study of the Nervous System and Behavior*, 44(8), 1105–32.
- Chong, S., & Whitelaw, E. (2004). Epigenetic germline inheritance. *Current Opinion in Genetics & Development*, 14(6), 692–6.
- Conway, C. M., & Christiansen, M. H. (2005). Modality-Constrained Statistical Learning of Tactile, Visual, and Auditory Sequences. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 31(1), 24–39.
- Conway, C. M., & Pisoni, D. B. (2008). Neurocognitive basis of implicit learning of sequential structure and its relation to language processing. *Annals of the New York Academy of Sciences*, 1145, 113–31.

- Creel, S. C., Newport, E. L., & Aslin, R. N. (2004). Distant melodies: statistical learning of nonadjacent dependencies in tone sequences. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, 30(5), 1119–30.
- Critten, S., Connelly, V., Dockrell, J. E., & Walter, K. (2014). Inflectional and derivational morphological spelling abilities of children with Specific Language Impairment. *Frontiers in Psychology*, 5, 948.
- Cunillera, T., Càmara, E., Toro, J. M., Marco-Pallares, J., Sebastián-Galles, N., Ortiz, H., ... Rodríguez-Fornells, A. (2009). Time course and functional neuroanatomy of speech segmentation in adults. *NeuroImage*, 48(3), 541–53.
- Cunillera, T., Toro, J. M., Sebastián-Gallés, N., & Rodríguez-Fornells, A. (2006). The effects of stress and statistical cues on continuous speech segmentation: an event-related brain potential study. *Brain Research*, 1123(1), 168–78.
- Daikoku, T., Yatomi, Y., & Yumoto, M. (2014). Implicit and explicit statistical learning of tone sequences across spectral shifts. *Neuropsychologia*, 63, 194–204.
- Dale, P. S., Dionne, G., Eley, T. C., & Plomin, R. (2000). Lexical and grammatical development: a behavioural genetic perspective. *Journal of Child Language*, 27(3), 619–42.
- Dale, P. S., Harlaar, N., Haworth, C. M. A., & Plomin, R. (2010). Two by two: a twin study of second-language acquisition. *Psychological Science*, 21(5), 635–40.
- Dale, P. S., Harlaar, N., Hayiou-Thomas, M. E., & Plomin, R. (2010). The etiology of diverse receptive language skills at 12 years. *Journal of Speech, Language, and Hearing Research : JSLHR*, 53(4), 982–92.
- De Diego Balaguer, R., Costa, A., Sebastián-Galles, N., Juncadella, M., & Caramazza, A. (2004). Regular and irregular morphology and its relationship with agrammatism: evidence from two Spanish-Catalan bilinguals. *Brain and Language*, 91(2), 212–22.
- De Diego Balaguer, R., Toro, J. M., Rodríguez-Fornells, A., & Bachoud-Lévi, A.-C. (2007). Different neurophysiological mechanisms

- underlying word and rule extraction from speech. *PloS One*, 2(11), e1175.
- De Geus, E. J. C., van't Ent, D., Wolfensberger, S. P. a, Heutink, P., Hoogendijk, W. J. G., Boomsma, D. I., & Veltman, D. J. (2007). Intrapair differences in hippocampal volume in monozygotic twins discordant for the risk for anxiety and depression. *Biological Psychiatry*, 61(9), 1062–71.
- De Moor, M. H. M., Boomsma, D. I., Stubbe, J. H., Willemsen, G., & de Geus, E. J. C. (2008). Testing causality in the association between regular exercise and symptoms of anxiety and depression. *Archives of General Psychiatry*, 65(8), 897–905.
- De Vries, M. H., Barth, A. C. R., Maiworm, S., Knecht, S., Zwitserlood, P., & Flöel, A. (2010). Electrical stimulation of Broca's area enhances implicit learning of an artificial grammar. *Journal of Cognitive Neuroscience*, 22(11), 2427–36.
- Deater-Deckard, K., Pike, A., Petrill, S. A., Cutting, A. L., Hughes, C., & O'Connor, T. G. (2001). Fast-Track Report: Nonshared environmental processes in social-emotional development: an observational study of identical twin differences in the preschool period. *Developmental Science*, 4(2), F1–F6.
- DeKeyser, R. (2000). The robustness of critical period effects in second language acquisition. *Studies in Second Language Acquisition*, 22(4), 499–533.
- Den Braber, A., van 't Ent, D., Boomsma, D. I., Cath, D. C., Veltman, D. J., Thompson, P. M., & de Geus, E. J. C. (2011). White matter differences in monozygotic twins discordant or concordant for obsessive-compulsive symptoms: a combined diffusion tensor imaging/voxel-based morphometry study. *Biological Psychiatry*, 70(10), 969–77.
- Den Braber, A., van 't Ent, D., Cath, D. C., Veltman, D. J., Boomsma, D. I., & de Geus, E. J. C. (2012). Brain activation during response interference in twins discordant or concordant for obsessive compulsive symptoms. *Twin Research and Human Genetics : The Official Journal of the International Society for Twin Studies*, 15(3), 372–83.

- Díaz, B., Baus, C., Escera, C., Costa, A., & Sebastián-Gallés, N. (2008). Brain potentials to native phoneme discrimination reveal the origin of individual differences in learning the sounds of a second language. *Proceedings of the National Academy of Sciences of the United States of America*, 105(42), 16083–8.
- Díaz, B., Sebastián-Gallés, N., Erdocia, K., Mueller, J. L., & Laka, I. (2011). On the cross-linguistic validity of electrophysiological correlates of morphosyntactic processing: A study of case and agreement violations in Basque. *Journal of Neurolinguistics*, 24(3), 357–373.
- Dick, A. S., & Tremblay, P. (2012). Beyond the arcuate fasciculus: consensus and controversy in the connectional anatomy of language. *Brain : A Journal of Neurology*, 135(Pt 12), 3529–50.
- Dionne, G., Dale, P. S., Boivin, M., & Plomin, R. (2003). Genetic Evidence for Bidirectional Effects of Early Lexical and Grammatical Development. *Child Development*, 74(2), 394–412.
- Domínguez, A., Alija, M., Cuetos, F., & de Vega, M. (2006). Event related potentials reveal differences between morphological (prefixes) and phonological (syllables) processing of words. *Neuroscience Letters*, 408(1), 10–5.
- Domínguez, A., de Vega, M., & Barber, H. (2004). Event-related brain potentials elicited by morphological, homographic, orthographic, and semantic priming. *Journal of Cognitive Neuroscience*, 16(4), 598–608.
- Doricchi, F., Thiebaut de Schotten, M., Tomaiuolo, F., & Bartolomeo, P. (2008). White matter (dis)connections and gray matter (dys)functions in visual neglect: gaining insights into the brain networks of spatial awareness. *Cortex; a Journal Devoted to the Study of the Nervous System and Behavior*, 44(8), 983–95.
- Duffau, H. (2006). Brain plasticity: from pathophysiological mechanisms to therapeutic applications. *Journal of Clinical Neuroscience*, 13(9), 885–897.
- Duffau, H. (2008). The anatomo-functional connectivity of language revisited. New insights provided by electrostimulation and tractography. *Neuropsychologia*, 46(4), 927–34.

- Eichenbaum, H. (2000). A cortical-hippocampal system for declarative memory. *Nature Reviews. Neuroscience*, 1(1), 41–50.
- Elman, J. (2006). Computational approaches to language acquisition. In *Encyclopedia of Language and Linguistics* (pp. 2–726).
- Elman, J. L. (2005). Connectionist models of cognitive development: where next? *Trends in Cognitive Sciences*, 9(3), 111–7.
- Endress, A. D. (2010). Learning melodies from non-adjacent tones. *Acta Psychologica*, 135(2), 182–90.
- Endress, A. D., Scholl, B. J., & Mehler, J. (2005). The Role of Salience in the Extraction of Algebraic Rules. *Journal of Experimental Psychology: General*, 134(3), 406–419.
- Epelbaum, S., Pinel, P., Gaillard, R., & Delmaire, C. (2008). Pure alexia as a disconnection syndrome: new diffusion imaging evidence for an old concept. *Cortex*, 44(8), 962–974.
- Evans, J. L., Saffran, J. R., & Robe-Torres, K. (2009). Statistical learning in children with specific language impairment. *Journal of Speech, Language, and Hearing Research : JSLHR*, 52(2), 321–35.
- Falconer, D., & Mackay, T. (1996). Heritability. In *Introduction to quantitative genetics* (4th ed., pp. 160–183). San Francisco: Benjamin Cummings.
- Finn, A. S., & Hudson Kam, C. L. (2008). The curse of knowledge: first language knowledge impairs adult learners' use of novel statistics for word segmentation. *Cognition*, 108(2), 477–99.
- Fischer, K. (1973). A comparison of the similarity in language skills of identical and fraternal twin pairs. Dissertation. *Abstracts International*, 34(8-a), 4864–4865.
- Fiser, J., & Aslin, R. N. (2001). Unsupervised statistical learning of higher-order spatial structures from visual scenes. *Psychological Science*, 12(6), 499–504.
- Fiser, J., & Aslin, R. N. (2002). Statistical learning of higher-order temporal structure from visual shape sequences. *Journal of*

- Experimental Psychology. Learning, Memory, and Cognition*, 28(3), 458–67.
- Flege, J. E., Bohn, O.-S., & Jang, S. (1997). Effects of experience on non-native speakers' production and perception of English vowels. *Journal of Phonetics*, 25(4), 437–470.
- Flege, J. E., & Liu, S. (2000). The Effect of Experience on Adults' Acquisition of A Second Language. *Studies in Second Language Acquisition*, 23(4), 527–52.
- Flöel, A., Vries, M. de, & Scholz, J. (2009). White matter integrity in the vicinity of Broca's area predicts grammar learning success. *Neuroimage*, 47(4), 1974–1981.
- Foch, T. T., & Plomin, R. (1980). Specific cognitive abilities in 5- to 12-year-old twins. *Behavior Genetics*, 10(6), 507–20.
- Forkstam, C., Hagoort, P., Fernandez, G., Ingvar, M., & Petersson, K. M. (2006). Neural correlates of artificial syntactic structure classification. *NeuroImage*, 32(2), 956–67.
- Fretham, S. J. B., Carlson, E. S., & Georgieff, M. K. (2011). The role of iron in learning and memory. *Advances in Nutrition (Bethesda, Md.)*, 2, 112–21.
- Friederici, A. D. (2009). Pathways to language: fiber tracts in the human brain. *Trends in Cognitive Sciences*, 13(4), 175–81.
- Friederici, A. D. (2012). The cortical language circuit: from auditory perception to sentence comprehension. *Trends in Cognitive Sciences*, 16(5), 262–8.
- Friederici, A. D., Steinhauer, K., & Pfeifer, E. (2002). Brain signatures of artificial language processing: evidence challenging the critical period hypothesis. *Proceedings of the National Academy of Sciences of the United States of America*, 99(1), 529–34.
- Gaffan, D., & Wilson, C. R. E. (2008). Medial temporal and prefrontal function: recent behavioural disconnection studies in the macaque monkey. *Cortex; a Journal Devoted to the Study of the Nervous System and Behavior*, 44(8), 928–35.

- Gathercole, S. E. (2006). Nonword repetition and word learning: The nature of the relationship. *Applied Psycholinguistics*, 27(04), 513–543.
- Gerner, G., & Baron, I. S. (2014). Pregnancy complications and neuropsychological outcomes: A review. *Child Neuropsychology: A Journal on Normal and Abnormal Development in Childhood and Adolescence*, 1–16.
- Geschwind, D. H., Miller, B. L., DeCarli, C., & Carmelli, D. (2002). Heritability of lobar brain volumes in twins supports genetic models of cerebral laterality and handedness. *Proceedings of the National Academy of Sciences of the United States of America*, 99(5), 3176–81.
- Gómez, R., & Gerken, L. (2000). Infant artificial language learning and language acquisition. *Trends in Cognitive Sciences*, 4(5), 178–186.
- Gómez, R. L. (2002). Variability and detection of invariant structure. *Psychological Science*, 13(5), 431–6.
- Gómez, R., & Maye, J. (2005). The Developmental Trajectory of Nonadjacent Dependency Learning. *Infancy*, 7(2), 183–206.
- Goodsitt, J. V, Morgan, J. L., & Kuhl, P. K. (1993). Perceptual strategies in prelingual speech segmentation. *Journal of Child Language*, 20(2), 229–52.
- Gopnik, M., & Goad, H. (1997). What underlies inflectional error patterns in genetic dysphasia? *Journal of Neurolinguistics*, 10(2-3), 109–137.
- Graf Estes, K., Evans, J. L., Alibali, M. W., & Saffran, J. R. (2007). Can infants map meaning to newly segmented words? Statistical segmentation and word learning. *Psychological Science*, 18(3), 254–60.
- Griffiths, T. L., & Kalish, M. L. (2007). Language evolution by iterated learning with bayesian agents. *Cognitive Science*, 31(3), 441–80.
- Grindstaff, J. L., Brodie, E. D., & Ketterson, E. D. (2003). Immune function across generations: integrating mechanism and evolutionary process in maternal antibody transmission. *Proceedings. Biological Sciences / The Royal Society*, 270(1531), 2309–19.

- Guerreiro, M., & Anguera, J. (2014). Age-equivalent Top–Down Modulation during Cross-modal Selective Attention. *Journal of Cognitive Neuroscience*, 26(12), 2827–2839.
- Häberling, I. S., Badzakova-Trajkov, G., & Corballis, M. C. (2013). Asymmetries of the arcuate fasciculus in monozygotic twins: genetic and nongenetic influences. *PloS One*, 8(1), e52315.
- Hart, B., & Risley, T. R. (1995). *Meaningful differences in the everyday experience of young American children*. (p. 268).
- Hart, B., & Risley, T. R. (1999). *The Social World of Children: Learning To Talk*. Paul H. Brookes Publishing Co., Baltimore, MD 21285-0624;
- Hasselquist, D., & Nilsson, J.-A. (2009). Maternal transfer of antibodies in vertebrates: trans-generational effects on offspring immunity. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, 364(1513), 51–60.
- Hauser, M. D., Newport, E. L., & Aslin, R. N. (2001). Segmentation of the speech stream in a non-human primate: statistical learning in cotton-top tamarins. *Cognition*, 78(3), B53–64.
- Haworth, C. M. A., Davis, O. S. P., & Plomin, R. (2013). Twins Early Development Study (TEDS): a genetically sensitive investigation of cognitive and behavioral development from childhood to young adulthood. *Twin Research and Human Genetics: The Official Journal of the International Society for Twin Studies*, 16(1), 117–25.
- Hayiou-Thomas, M. E., Dale, P. S., & Plomin, R. (2012). The etiology of variation in language skills changes with development: a longitudinal twin study of language from 2 to 12 years. *Developmental Science*, 15(2), 233–49.
- Hayiou-Thomas, M. E., Harlaar, N., Dale, P. S., & Plomin, R. (2006). Genetic and environmental mediation of the prediction from preschool language and nonverbal ability to 7-year reading. *Journal of Research in Reading*, 29(1), 50–74.
- Hayiou-Thomas, M. E., Harlaar, N., Dale, P. S., & Plomin, R. (2010). Preschool speech, language skills, and reading at 7, 9, and 10 years:

- etiology of the relationship. *Journal of Speech, Language, and Hearing Research : JSLHR*, 53(2), 311–32.
- Hensch, T. K. (2005). Critical period plasticity in local cortical circuits. *Nature Reviews. Neuroscience*, 6(11), 877–88.
- Hickok, G., & Poeppel, D. (2007). The cortical organization of speech processing. *Nature Reviews Neuroscience*, 8, 393–402.
- Hoekstra, R. a, Bartels, M., van Leeuwen, M., & Boomsma, D. I. (2009). Genetic architecture of verbal abilities in children and adolescents. *Developmental Science*, 12(6), 1041–53.
- Hoff, E. (2003). The Specificity of Environmental Influence: Socioeconomic Status Affects Early Vocabulary Development Via Maternal Speech. *Child Development*, 74(5), 1368–1378.
- Hoff, E. (2006). How social contexts support and shape language development. *Developmental Review*, 26(1), 55–88.
- Hohnen, B., & Stevenson, J. (1999). The structure of genetic influences on general cognitive, language, phonological, and reading abilities. *Developmental Psychology*, 35(2), 590–603.
- Hopp, H., & Schmid, M. S. (2011). Perceived foreign accent in first language attrition and second language acquisition: The impact of age of acquisition and bilingualism. *Applied Psycholinguistics*, 34(02), 361–394.
- Hornickel, J., Lin, D., & Kraus, N. (2013). Speech-evoked auditory brainstem responses reflect familial and cognitive influences. *Developmental Science*, 16(1), 101–10.
- Hsu, H. J., Tomblin, J. B., & Christiansen, M. H. (2014). Impaired statistical learning of non-adjacent dependencies in adolescents with specific language impairment. *Frontiers in Psychology*, 5, 175.
- Hudon, T. M., Fennell, C. T., & Hoftyzer, M. (2013). Quality not quantity of television viewing is associated with bilingual toddlers' vocabulary scores. *Infant Behavior & Development*, 36(2), 245–54.

- Huttenlocher, P. R. (1999). Dendritic and Synaptic Development in Human Cerebral Cortex: Time Course and Critical Periods. *Developmental Neuropsychology*, 16(3), 347–349.
- Huttenlocher, P. R., & Dabholkar, A. S. (1997). Regional differences in synaptogenesis in human cerebral cortex. *The Journal of Comparative Neurology*, 387(2), 167–78.
- Jakoby, H., Goldstein, A., & Faust, M. (2011). Electrophysiological correlates of speech perception mechanisms and individual differences in second language attainment. *Psychophysiology*, 48(11), 1517–31.
- Jedynak, M. (2009). *Critical Period Hypothesis Revisited: The Impact of Age on Ultimate Attainment in the Pronunciation of a Foreign Language* (p. 263). Peter Lang.
- Jin, Y., Díaz, B., Colomer, M., & Sebastián-Gallés, N. (2014). Oscillation encoding of individual differences in speech perception. *PloS One*, 9(7), e100901.
- Johnson, J., & Newport, E. (1989). Critical period effects in second language learning: The influence of maturational state on the acquisition of English as a second language. *Cognitive Psychology*, 21, 60–99.
- Johnson, S. P., Fernandes, K. J., Frank, M. C., Kirkham, N., Marcus, G., Rabagliati, H., & Slemmer, J. A. (2009). Abstract Rule Learning for Visual Sequences in 8- and 11-Month-Olds. *Infancy: The Official Journal of the International Society on Infant Studies*, 14(1), 2–18.
- Kang, C., & Drayna, D. (2011). Genetics of speech and language disorders 1. *Annual Review of Genomics and Human Genetics*, 12, 145–64.
- Karuza, E. a, Newport, E. L., Aslin, R. N., Starling, S. J., Tivarus, M. E., & Bavelier, D. (2013). The neural correlates of statistical learning in a word segmentation task: An fMRI study. *Brain and Language*, 127(1), 46–54.
- Kirkham, N. Z., Slemmer, J. a, & Johnson, S. P. (2002). Visual statistical learning in infancy: evidence for a domain general learning mechanism. *Cognition*, 83(2), B35–42.

- Kovas, Y., Hayiou-Thomas, M. E., Oliver, B., Dale, P. S., Bishop, D. V. M., & Plomin, R. (2005). Genetic influences in different aspects of language development: the etiology of language skills in 4.5-year-old twins. *Child Development*, 76(3), 632–51.
- Kovas, Y., & Hayiou-Thomas, M. (2005). Genetic Influences in Different Aspects of Language Development: The Etiology of Language Skills in 4.5-Year-Old Twins. *Child Development*, 76(3), 632–551.
- Kuhl, P. K., Tsao, F.-M., & Liu, H.-M. (2003). Foreign-language experience in infancy: effects of short-term exposure and social interaction on phonetic learning. *Proceedings of the National Academy of Sciences of the United States of America*, 100(15), 9096–101.
- Langus, A., Marchetto, E., Bion, R. A. H., & Nespor, M. (2012). Can prosody be used to discover hierarchical structure in continuous speech? *Journal of Memory and Language*, 66(1), 285–306.
- Lelekov, T., Dominey, P. F., & Garcia-Larrea, L. (2000). Dissociable ERP profiles for processing rules vs instances in a cognitive sequencing task. *Neuroreport*, 11(5), 1129–32.
- Leonard, L., & Eyer, J. (1996). Deficits of grammatical morphology in children with specific language impairment and their implications for notions of bootstrapping. In *Signal to syntax. bootstrapping from speech to grammar in early acquisition* (pp. 233–247).
- Leroy, F., Glasel, H., Dubois, J., Hertz-Pannier, L., Thirion, B., Mangin, J.-F., & Dehaene-Lambertz, G. (2011). Early maturation of the linguistic dorsal pathway in human infants. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, 31(4), 1500–6.
- Liang, H., & Eley, T. C. (2005). A monozygotic twin differences study of nonshared environmental influence on adolescent depressive symptoms. *Child Development*, 76(6), 1247–60.
- Liao, D. A., Kronemer, S. I., Yau, J. M., Desmond, J. E., & Marvel, C. L. (2014). Motor system contributions to verbal and non-verbal working memory. *Frontiers in Human Neuroscience*, 8, 753.

- Linck, J. A., Kroll, J. F., & Sunderman, G. (2009). Losing access to the native language while immersed in a second language: evidence for the role of inhibition in second-language learning. *Psychological Science*, 20(12), 1507–15.
- López-Barroso, D., Catani, M., Ripollés, P., Dell’Acqua, F., Rodríguez-Fornells, A., & de Diego-Balaguer, R. (2013). Word learning is mediated by the left arcuate fasciculus. *Proceedings of the National Academy of Sciences of the United States of America*, 110(32), 13168–73.
- Lopez-Barroso, D., de Diego-Balaguer, R., Cunillera, T., Camara, E., Münte, T. F., & Rodríguez-Fornells, A. (2011). Language learning under working memory constraints correlates with microstructural differences in the ventral language pathway. *Cerebral Cortex (New York, N.Y. : 1991)*, 21(12), 2742–50.
- Marchman, V. A., Martinez-Sussmann, C., & Dale, P. S. (2004). The language-specific nature of grammatical development: evidence from bilingual language learners. *Developmental Science*, 7(2), 212–224.
- Marcus, G. F., Fernandes, K. J., & Johnson, S. P. (2007). Infant rule learning facilitated by speech. *Psychological Science*, 18(5), 387–91.
- Marcus, G. F., Vijayan, S., Bandi Rao, S., & Vishton, P. M. (1999). Rule learning by seven-month-old infants. *Science (New York, N.Y.)*, 283(5398), 77–80.
- Mather, P. L., & Black, K. N. (1983). Hereditary and Environmental Influences on Preschool Twins’ Language Skills. *Developmental Psychology*, 20(2), 303–8.
- Maye, J., Weiss, D. J., & Aslin, R. N. (2008). Statistical phonetic learning in infants: facilitation and feature generalization. *Developmental Science*, 11(1), 122–34.
- Maye, J., Werker, J. F., & Gerken, L. (2002). Infant sensitivity to distributional information can affect phonetic discrimination. *Cognition*, 82(3), B101–11.

- McDonald, C., Ahmadi, M., & Hagler, D. (2008). Diffusion tensor imaging correlates of memory and language impairments in temporal lobe epilepsy. *Neurology*, 71(23), 1869–1876.
- McDonald, C. R., Leyden, K. M., Hagler, D. J., Kucukboyaci, N. E., Kemmotsu, N., Tecoma, E. S., & Iragui, V. J. (2014). White matter microstructure complements morphometry for predicting verbal memory in epilepsy. *Cortex; a Journal Devoted to the Study of the Nervous System and Behavior*, 58, 139–150.
- McNealy, K., Mazziotta, J. C., & Dapretto, M. (2006). Cracking the language code: neural mechanisms underlying speech parsing. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, 26(29), 7629–39.
- Miller, G. A., & Selfridge, J. A. (1950). Verbal context and the recall of meaningful material. *The American Journal of Psychology*, 63(2), 176–85.
- Mirman, D., Magnuson, J. S., Estes, K. G., & Dixon, J. A. (2008). The link between statistical segmentation and word learning in adults. *Cognition*, 108(1), 271–80.
- Mueller, J. L., Bahlmann, J., & Friederici, A. D. (2008). The role of pause cues in language learning: the emergence of event-related potentials related to sequence processing. *Journal of Cognitive Neuroscience*, 20(5), 892–905.
- Mueller, J. L., Hahne, A., Fujii, Y., & Friederici, A. D. (2005). Native and nonnative speakers' processing of a miniature version of Japanese as revealed by ERPs. *Journal of Cognitive Neuroscience*, 17(8), 1229–44.
- Murphy, R. A., Mondragón, E., & Murphy, V. A. (2008). Rule learning by rats. *Science (New York, N.Y.)*, 319(5871), 1849–51.
- Nash, M., & Donaldson, M. L. (2005). Word learning in children with vocabulary deficits. *Journal of Speech, Language, and Hearing Research: JSLHR*, 48(2), 439–58.
- Nelson, C. A., de Haan, M., & Thomas, K. M. (2012). *Neuroscience of cognitive development: The role of experience and the developing brain*. (p. 215). John Wiley & Sons.

- Newbury, D. F., & Monaco, a P. (2010). Genetic advances in the study of speech and language disorders. *Neuron*, 68(2), 309–20.
- Newport, E. L., Hauser, M. D., Spaepen, G., & Aslin, R. N. (2004). Learning at a distance II. Statistical learning of non-adjacent dependencies in a non-human primate. *Cognitive Psychology*, 49(2), 85–117.
- Olson, L. L., & Samuels, S. J. (1973). The relationship between age and accuracy of foreign language pronunciation. *The Journal of Educational Research*, 66(263-268).
- Opitz, B., & Friederici, A. D. (2003). Interactions of the hippocampal system and the prefrontal cortex in learning language-like rules. *NeuroImage*, 19(4), 1730–7.
- Opitz, B., & Friederici, A. D. (2004). Brain correlates of language learning: the neuronal dissociation of rule-based versus similarity-based learning. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, 24(39), 8436–40.
- Ortega, L. (2014). *Understanding Second Language Acquisition*. *Understanding second language acquisition* (p. 296). Routledge.
- Pelucchi, B., Hay, J. F., & Saffran, J. R. (2009). Statistical learning in a natural language by 8-month-old infants. *Child Development*, 80(3), 674–85.
- Peña, M., Bonatti, L. L., Nespor, M., & Mehler, J. (2002). Signal-driven computations in speech processing. *Science (New York, N.Y.)*, 298(5593), 604–7.
- Perruchet, P., & Pacton, S. (2006). Implicit learning and statistical learning: one phenomenon, two approaches. *Trends in Cognitive Sciences*, 10(5), 233–8.
- Perruchet, P., Peereman, R., & Tyler, M. D. (2006). Do we need algebraic-like computations? A reply to Bonatti, Peña, Nespor, and Mehler (2006). *Journal of Experimental Psychology: General*, 135(2), 322–326.

- Perruchet, P., Tyler, M. D., Galland, N., & Peereman, R. (2004). Learning nonadjacent dependencies: no need for algebraic-like computations. *Journal of Experimental Psychology. General*, 133(4), 573–83.
- Petersson, K. (2004). Artificial syntactic violations activate Broca's region. *Cognitive Science*, 28(3), 383–407.
- Petrides, M., Alivisatos, B., Meyer, E., & Evans, A. C. (1993). Functional activation of the human frontal cortex during the performance of verbal working memory tasks. *Proceedings of the National Academy of Sciences*, 90(3), 878–882.
- Pike, A., Reiss, D., Hetherington, E. M., & Plomin, R. (1996). Using MZ differences in the search for nonshared environmental effects. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, 37(6), 695–704.
- Plomin, R., DeFries, J. C., Knopik, V. S., & Neiderhiser, J. M. (2013). *Behavioral Genetics* (6th ed., p. 560). Worth Publishers.
- Poeppel, D., & Hickok, G. (2004). Dorsal and ventral streams: a framework for understanding aspects of the functional anatomy of language. *Cognition*, 92(1-2), 1–12.
- Posner, M. I., Sheese, B. E., Odludas, Y., & Tang, Y. (2006). Analyzing and shaping human attentional networks. *Neural Networks: The Official Journal of the International Neural Network Society*, 19(9), 1422–9.
- Pujol, J., Soriano-Mas, C., Ortiz, H., Sebastián-Gallés, N., Losilla, J. M., & Deus, J. (2006). Myelination of language-related areas in the developing brain. *Neurology*, 66(3), 339–43.
- Rassoulzadegan, M., Grandjean, V., Gounon, P., Vincent, S., Gillot, I., & Cuzin, F. (2006). RNA-mediated non-mendelian inheritance of an epigenetic change in the mouse. *Nature*, 441(7092), 469–74.
- Reber, A. S. (1967). Implicit learning of artificial grammars. *Journal of Verbal Learning and Verbal Behavior*, 6(6), 855–863.
- Reber, P. J. (2013). The neural basis of implicit learning and memory: a review of neuropsychological and neuroimaging research. *Neuropsychologia*, 51(10), 2026–42.

- Reinke, K. S., He, Y., Wang, C., & Alain, C. (2003). Perceptual learning modulates sensory evoked response during vowel segregation. *Brain Research. Cognitive Brain Research*, 17(3), 781–91.
- Rice, M. L., Haney, K. R., & Wexler, K. (1998). Family histories of children with SLI who show extended optional infinitives. *Journal of Speech, Language, and Hearing Research : JSLHR*, 41(2), 419–32.
- Roberts, J. E., Burchinal, M., & Durham, M. (1999). Parents' Report of Vocabulary and Grammatical Development of African American Preschoolers: Child and Environmental Associations. *Child Development*, 70(1), 92–106.
- Ross, E. D. (2008). Sensory-specific amnesia and hypoemotionality in humans and monkeys: gateway for developing a hodology of memory. *Cortex; a Journal Devoted to the Study of the Nervous System and Behavior*, 44(8), 1010–22.
- Saffran, J., Newport, E. L., & Aslin, R. (1996). Word segmentation: The role of distributional cues. *Journal of Memory and Language*, 62(35), 606–621.
- Saffran, J. R. (2001). Words in a sea of sounds: the output of infant statistical learning. *Cognition*, 81(2), 149–69.
- Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996). Statistical learning by 8-month-old infants. *Science*, 274(5294), 1926–1928.
- Saffran, J. R., Johnson, E. K., Aslin, R. N., & Newport, E. L. (1999). Statistical learning of tone sequences by human infants and adults. *Cognition*, 70(1), 27–52.
- Saffran, J. R., Newport, E. L., Aslin, R., Tunick, R. A., & Barrueco, S. (1997). Incidental language learning: Listening (and Learning) out of the Comer of Your Ear. *Psychological Science*, 8(2), 101–105.
- Saffran, J. R., Pollak, S. D., Seibel, R. L., & Shkolnik, A. (2007). Dog is a dog is a dog: infant rule learning is not specific to language. *Cognition*, 105(3), 669–80.
- Sakai, K. L., Miura, K., Narafu, N., & Muraishi, Y. (2004). Correlated functional changes of the prefrontal cortex in twins induced by

- classroom education of second language. *Cerebral Cortex (New York, N.Y. : 1991)*, 14(11), 1233–9.
- Sanders, L. D., Newport, E. L., & Neville, H. J. (2002). Segmenting nonsense: an event-related potential index of perceived onsets in continuous speech. *Nature Neuroscience*, 5(7), 700–3. doi:10.1038/nn873
- Sandoval, M., & Gómez, R. L. (2013). The development of nonadjacent dependency learning in natural and artificial languages. *Wiley Interdisciplinary Reviews: Cognitive Science*, 4(5), 511–522.
- Saur, D., Kreher, B. W., Schnell, S., Kümmerer, D., Kellmeyer, P., Vry, M.-S., ... Weiller, C. (2008). Ventral and dorsal pathways for language. *Proceedings of the National Academy of Sciences of the United States of America*, 105(46), 18035–40.
- Schapiro, A. C., Gregory, E., Landau, B., McCloskey, M., & Turk-Browne, N. B. (2014). The necessity of the medial temporal lobe for statistical learning. *Journal of Cognitive Neuroscience*, 26(8), 1736–47.
- Sebastián-Gallés, N., & Baus, C. (2005). On the relationship between perception and production in L2 categories. In A. Cutler (Ed.), *Twenty-first century Psycholinguistics: Four cornerstones* (pp. 279–292).
- Sebastián-Gallés, N., & Díaz, B. (2012). First and Second Language Speech Perception: Graded Learning. *Language Learning*, 62, 131–147.
- Sebastián-Gallés, N., Soriano-Mas, C., Baus, C., Díaz, B., Ressel, V., Pallier, C., ... Pujol, J. (2012). Neuroanatomical markers of individual differences in native and non-native vowel perception. *Journal of Neurolinguistics*, 25(3), 150–162.
- Segal, N. L. (1984). Monozygotic and Dizygotic Twins: A Comparative Analysis of Mental Ability Profiles. *Child Development*, 56(4), 1051–58.
- Seidenberg, M. S. (1997). Language Acquisition and Use: Learning and Applying Probabilistic Constraints. *Science*, 275(5306), 1599–1603.

- Shapiro, K. A., Moo, L. R., & Caramazza, A. (2012). Neural Specificity for Grammatical Operations is Revealed by Content-Independent fMR Adaptation. *Frontiers in Psychology*, 3, 26.
- Shorter, J., & Lindquist, S. (2005). Prions as adaptive conduits of memory and inheritance. *Nature Reviews. Genetics*, 6(6), 435–50.
- Slevc, L. R., & Miyake, A. (2006). Individual differences in second-language proficiency: does musical ability matter? *Psychological Science*, 17(8), 675–81.
- SLI Consortium. (2002). A genomewide scan identifies two novel loci involved in specific language impairment. *American Journal of Human Genetics*, 70(2), 384–98.
- SLI Consortium. (2004). Highly significant linkage to the SLI1 locus in an expanded sample of individuals affected by specific language impairment. *American Journal of Human Genetics*, 74(6), 1225–38.
- Sowell, E. R., Thompson, P. M., Leonard, C. M., Welcome, S. E., Kan, E., & Toga, A. W. (2004). Longitudinal mapping of cortical thickness and brain growth in normal children. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, 24(38), 8223–31.
- Stevens, D. J., Arciuli, J., & Anderson, D. I. (2014). Concurrent Movement Impairs Incidental But Not Intentional Statistical Learning. *Cognitive Science*.
- Stromswold, K. (2001). The Heritability of Language: A Review and Metaanalysis of Twin, Adoption, and Linkage Studies. *Language*, 77(4), 647–723.
- Thompson, L. A., Detterman, D. K., & Plomin, R. (1991). Associations between cognitive abilities and scholastic achievement: genetic overlap but environmental differences. *Psychological Science*, 2(3), 158–165.
- Thompson, S. P., & Newport, E. L. (2007). Statistical Learning of Syntax: The Role of Transitional Probability. *Language Learning and Development*, 3(1), 1–42.

- Toro, J. M., Pons, F., Bion, R. A. H., & Sebastián-Gallés, N. (2011). The contribution of language-specific knowledge in the selection of statistically-coherent word candidates. *Journal of Memory and Language*, 64(2), 171–180.
- Toro, J. M., Trobalon, J. B., & Sebastián-Gallés, N. (2005). Effects of Backward Speech and Speaker Variability in Language Discrimination by Rats. *Journal of Experimental Psychology: Animal Behavior Processes*, 31(1), 95–100.
- Tsukada, K., Birdsong, D., Bialystok, E., Mack, M., Sung, H., & Flege, J. (2005). A developmental study of English vowel production and perception by native Korean adults and children. *Journal of Phonetics*, 33(3), 263–290.
- Turk-Browne, N. B., Jungé, J., & Scholl, B. J. (2005). The automaticity of visual statistical learning. *Journal of Experimental Psychology: General*, 134(4), 552–64.
- Turkheimer, E., & Waldron, M. (2000). Nonshared environment: a theoretical, methodological, and quantitative review. *Psychological Bulletin*, 126(1), 78–108.
- Uddén, J., Folia, V., Forkstam, C., Ingvar, M., Fernandez, G., Overeem, S., ... Petersson, K. M. (2008). The inferior frontal cortex in artificial syntax processing: an rTMS study. *Brain Research*, 1224, 69–78.
- Ullman, M. T. (2001). The Neural Basis of Lexicon and Grammar in First and Second Language: The Declarative/Procedural Model. *Bilingualism: Language and Cognition*, 4(2), 105–22.
- Ullman, M. T. (2004). Contributions of memory circuits to language: the declarative/procedural model. *Cognition*, 92(1-2), 231–70.
- Ullman, M. T. (2005). A cognitive neuroscience perspective on second language acquisition: The declarative/procedural model. In *Mind and Context in Adult Second Language Acquisition: Methods, Theory, and Practice* (pp. 141–178).
- Van 't Ent, D., Lehn, H., Derks, E. M., Hudziak, J. J., Van Strien, N. M., Veltman, D. J., ... Boomsma, D. I. (2007). A structural MRI study in monozygotic twins concordant or discordant for

- attention/hyperactivity problems: evidence for genetic and environmental heterogeneity in the developing brain. *NeuroImage*, 35(3), 1004–20.
- Van Dongen, J., Slagboom, P. E., Draisma, H. H. M., Martin, N. G., & Boomsma, D. I. (2012). The continuing value of twin studies in the omics era. *Nature Reviews. Genetics*, 13(9), 640–53.
- Van Soelen, I. L. C., van den Berg, S. M., Dekker, P. H., van Leeuwen, M., Peper, J. S., Hulshoff Pol, H. E., & Boomsma, D. I. (2009). Individual differences in dynamic measures of verbal learning abilities in young twin pairs and their older siblings. *Learning and Individual Differences*, 19(4), 440–444.
- Visscher, P. M. (2004). Power of the classical twin design revisited. *Twin Research : The Official Journal of the International Society for Twin Studies*, 7(5), 505–12.
- Walsh, M., Montojo, C. A., Sheu, Y.-S., Marchette, S. A., Harrison, D. M., Newsome, S. D., ... Courtney, S. M. (2011). Object working memory performance depends on microstructure of the frontal-occipital fasciculus. *Brain Connectivity*, 1(4), 317–29.
- Wang, T., & Saffran, J. R. (2014). Statistical learning of a tonal language: the influence of bilingualism and previous linguistic experience. *Frontiers in Psychology*, 5, 953.
- Weber-Fox, C. M., & Neville, H. J. (1996). Maturational Constraints on Functional Specializations for Language Processing: ERP and Behavioral Evidence in Bilingual Speakers. *Journal of Cognitive Neuroscience*, 8(3), 231–56.
- Wong, P. C. M., Morgan-Short, K., Ettlinger, M., & Zheng, J. (2012). Linking neurogenetics and individual differences in language learning: the dopamine hypothesis. *Cortex; a Journal Devoted to the Study of the Nervous System and Behavior*, 48(9), 1091–102.
- Wong, P. C. M., Perrachione, T. K., & Parrish, T. B. (2007). Neural characteristics of successful and less successful speech and word learning in adults. *Human Brain Mapping*, 28(10), 995–1006.

- Wright, A. A., & Rivera, J. J. (1997). Memory of auditory lists by rhesus monkeys (*Macaca mulatta*). *Journal of Experimental Psychology. Animal Behavior Processes*, 23(4), 441–9.
- Yamagata, S., Takahashi, Y., Ozaki, K., Fujisawa, K. K., Nonaka, K., & Ando, J. (2013). Bidirectional influences between maternal parenting and children's peer problems: a longitudinal monozygotic twin difference study. *Developmental Science*, 16(2), 249–59.
- Young, J. P., Lader, M. H., & Fenton, G. W. (1972). A twin study of the genetic influences on the electroencephalogram. *Journal of Medical Genetics*, 9(1), 13–6.
- Zendel, B. R., & Alain, C. (2012). Musicians experience less age-related decline in central auditory processing. *Psychology and Aging*, 27(2), 410–417.

ANNEX

A summary of an additional and ongoing study is presented in this annex.

Exploring the genetic and environmental influence in both native and non-native phoneme perception

1. Introduction

Learning the sounds of a second language (L2) is a hard challenge that not everyone achieves with the same ease. It has been shown that the most relevant factors contributing to individual differences on L2 perception are the age in which the L2 is acquired as well as the time of exposure to this new language. Indeed, previous studies have revealed that learners who have been exposed to an L2 during childhood and/or in an immersed environment generally attain a better command of L2 perception than those who have been exposed late and/or have received less exposure (Bialystok & Hakuta, 1999; DeKeyser, 2000; Flege et al., 1997; Johnson & Newport, 1989; Weber-Fox & Neville, 1996).

The study of the neural substrates causing variation in L2 phoneme perception have provided a more clear idea of the origin of individual differences (Díaz et al., 2008; Golestani & Pallier, 2007; Golestani & Zatorre, 2009; Jakoby et al., 2011; Jin et al., 2014;

Sebastián-Gallés & Díaz, 2012). For instance, Díaz et al. (2008) provided a striking clue on this issue. They tested good and bad L2 perceivers (all participants were early learners) in their ability to process both tones varying in pitch and duration and native and non-native phonetic contrasts. By measuring participants' mismatch negativity (MMN) they only observed differences between groups in the speech conditions and differences were more prominent in the native condition. Hence, from these findings two conclusions were reached: first, they suggested that being a good or bad L2 perceiver would not be related with general-acoustic abilities but language-specific ones and, second, variation in L2 perception would be at the core of first language (L1) perceptual skills. This last hypothesis would be in agreement with those suggesting the share of neural underpinnings between L1 and L2 processing (Golestani & Zatorre, 2004; Perani & Abutalebi, 2005).

Importantly, participants in Díaz et al. (2008) were comparable in terms of age of acquisition and time of exposure to an L2, which suggests that differences between them were not caused by these two factors. Hence, it seems that our innate predisposition (e.g., genetic endowment) also might play a role on the L2 perception ability. Unfortunately, little research has focused on the role of genetics in L2 learning. In fact, the only study investigating heritability (i.e., the percentage of genetic variance explaining the total variance of a phenotype) in second language is Dale et al. (2010). They evaluated L2 learning in 14-years-old twins by collecting teachers' assessments. However, they did not calculate

heritability for specific tasks, thus making impossible to know the influence of genetics in phoneme perception. Other studies focused on L1 learning in children found that genetic variance for tasks related with phoneme perception were between 40 and 50% (Bishop et al., 2006; Byrne et al., 2013; Hayiou-Thomas, Harlaar, Dale, & Plomin, 2006; Kovas & Hayiou-Thomas, 2005). Hornickel, Lin, & Kraus (2013) were interested on studying the impact of shared genetic and environmental factors in the auditory brainstem response to speech. Although they did not measure heritability, they calculated the similarity between pair of children with different levels of relatedness (i.e., siblings, reading-matched and age/sex-matched). The results showed a graded pattern of similarity according to the level of relatedness, which suggested that sibling-ship (the share of environment and/or genetics) and auditory-based communication skills such as reading would have an impact on speech perception. With a similar purpose as Hornickel et al. (2013), here we aim at studying the impact of genetic and environmental sources on the perception of both native and non-native phonemes.

2. Methods

We selected a group of 68 MZT (34 pairs) and 56 DZT (28 pairs) and we used the procedure followed by Díaz et al. (2008). Participants were presented with two speech sound contrasts conditions following an oddball paradigm (i.e., Native condition (N): deviant /e/ and standard /o/, Non-native condition (NN): deviant /ö/ and standard /o/) while brain signal was recorded. First, we performed an event-related potential (ERP) analysis in which both

zygosity and sibling-ship was controlled in order to replicate Díaz et al. (2008) findings. Second, a wavelet analysis was applied and the intra-pair cross-correlation was measured in order to know how similar was the neural activity for each pair of twins in the time window corresponding to the MMN (Hornickel et al., 2013; Young, Lader, & Fenton, 1972). We hypothesized that, if genetics have a significant role on native and non-native phoneme perception as previous studies suggested, we may observe a higher similarity for MZT than DZT. Moreover, if both native and non-native phoneme perception are governed by the same neural mechanisms, we may not observed differences in the implication of both genetic and environmental factors. However, if it is the case, we cannot discard other possible hypotheses such as the implication of different mechanisms with the same genetic/environmental load.

3. Results

Results from the ERP analysis showed that, as Díaz et al. (2008), a reliable MMN was elicited for both native and non-native condition in both groups. That is, the amplitude of the MMN at the central and frontal electrodes (Cz, Fz) was compared against zero level separately for each group of participants (*see Table 1*).

A linear mixed effects regression analysis was applied in order to compare the mean activity at the Fz and Cz electrodes between groups (MZT and DZT) and conditions (N and NN). As expected, differences between conditions were observed ($\beta=-0.46$, S.E.=0.22, $t_{(407)}=-2.10$, $p<0.04$) but not for groups ($\beta=-0.18$, S.E.=0.29,

$t_{(59)}=0.64$, $p=0.53$). No interaction was significant either (All $p>0.09$). Hence, the significant differences observed between native and non-native conditions replicated previous findings (Díaz et al., 2008) showing larger MMN amplitude for the N than for the NN condition (*see figure 1*).

Condition	Group	Electrode	Latency	Mean Ampl.	<i>t</i>	d.f.
N	MZT	Fz	115-145	-1.59	-7.09**	67
		Cz	100-140	-1.88	-8.34**	67
	DZT	Fz	122-162	-1.58	-8.41**	55
		Cz	114-154	-1.96	-8.34**	55
NN	MZT	Fz	137-177	-1.26	-5.36**	67
		Cz	122-162	-1.01	-4.18**	67
	DZT	Fz	149-189	-1.35	-4.91**	55
		Cz	132-172	-1.27	-4.62**	55

Table A1. *T* tests against 0 level of the MMN mean amplitude for the native phoneme and nonnative phoneme. Significant differences: **, $p<0.001$; d.f., degrees of freedom.

A continuous wavelet analysis at the Cz electrode was carried out for each subject in the signal elicited by the deviant stimuli of each condition (N: /e/; NN: /ö/). Once obtained the wavelet map, a cross-correlation analysis was applied to measure the similarity between twins in the time window corresponding to the appearance of the MMN. The mean of the maximum cross-correlation for the native condition was $r=0.65$ and $r=0.61$ for MZT and DZT, respectively. In the non-native condition the maximum cross-correlation was $r=0.66$ for MZT and $r=0.63$ for DZT (*see figure 2*). Afterwards, an analysis of variance was performed in which both N and NN conditions were

compared between groups. Results showed neither significant main effects nor an interaction (all $F_s < 1$).

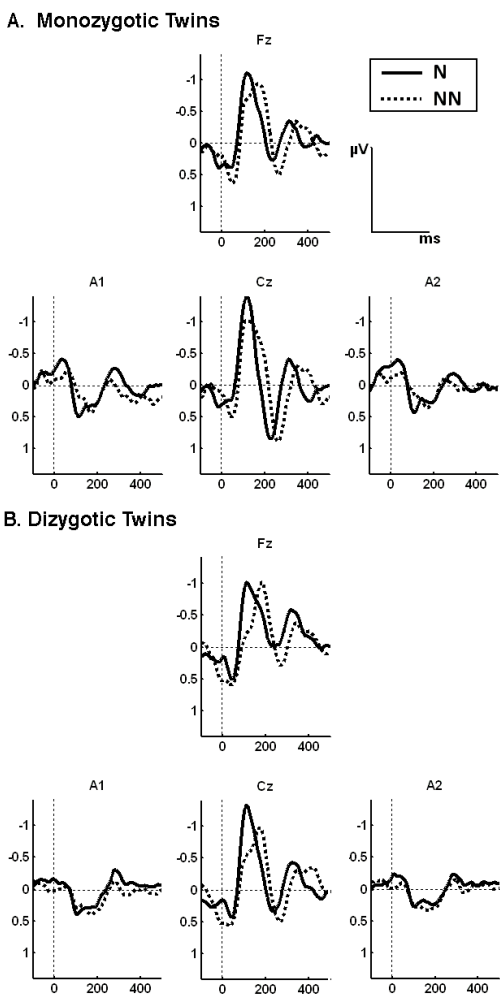


Figure A1. ERP grand average at Fz, Cz and mastoid electrodes for the native (full line) and non-native (dotted line) condition in (A) MZT and (B) DZT. A. In both groups the native condition elicited a larger MMN.

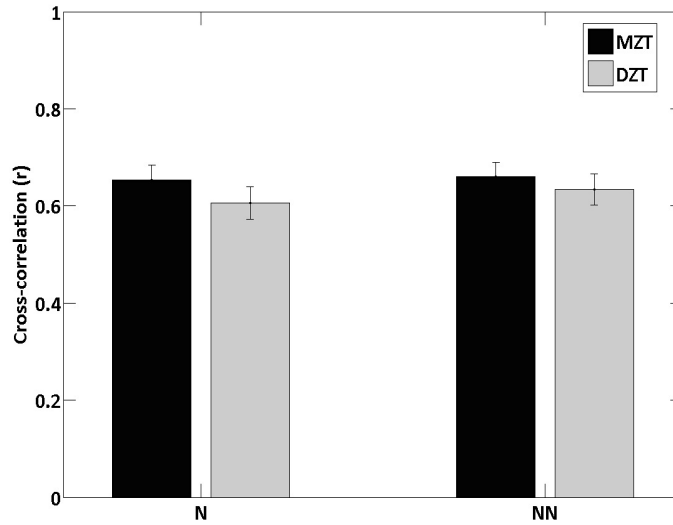


Figure A2. Cross-correlation (r) mean value for MZT and DZT in each condition (N and NN). Error bars indicate SE.

5. Discussion

The main goal of this study was to investigate the involvement of genetic and environmental pressures on the ability to perceive native and non-native phonemes. The study was divided in two phases. In the first phase, we performed an ERP analysis in order to know whether the procedure followed elicited similar electrophysiological responses to the ones reported in Díaz et al. (2008). From this analysis we observed that the perception of native phoneme elicited a larger and earlier MMN in both MZT and DZT compared to the non-native phoneme. Differences between groups were not found. Hence, from this first phase we conclude that both MZT and DZT were sensitive to the differences between the native and non-native phonetic contrasts.

In the second phase, we compare the brain activity between twin pairs in the deviant stimuli (/e/ and /ö/) by applying a wavelet analysis and measuring the cross-correlation in the time window corresponding to the appearance of the MMN. Results showed neither differences between groups nor between native and non-native conditions.

Being speculative, the results of this second phase would suggest that the perception of native and non-native phonemes would be mainly explained by environmental factors. Moreover, the lack of differences between phonemes would suggest that the same neural devices might carry their processing. However, further research is necessary in order to confirm these hypotheses.

References

- Bialystok, E., & Hakuta, K. (1999). *Confounded age: Linguistic and cognitive factors in age differences for second language acquisition*. (T. & F. E-Library, Ed.) *Second language acquisition and the critical period hypothesis*. Lawrence Erlbaum Associates, Inc.
- Bishop, D. V. M., Adams, C. V., & Norbury, C. F. (2006). Distinct genetic influences on grammar and phonological short-term memory deficits: evidence from 6-year-old twins. *Genes, Brain, and Behavior*, 5(2), 158–69.
- Byrne, B., Olson, R. K., Samuelsson, S., Wadsworth, S., Corley, R., DeFries, J. C., & Willcutt, E. (2013). Genetic and environmental influences on early literacy. *Journal of Research in Reading*, 29(1), 33–49.
- DeKeyser, R. (2000). The robustness of critical period effects in second language acquisition. *Studies in Second Language Acquisition*, 22(4), 499–533.
- Díaz, B., Baus, C., Escera, C., Costa, A., & Sebastián-Gallés, N. (2008). Brain potentials to native phoneme discrimination reveal the origin of individual differences in learning the sounds of a second language. *Proceedings of the National Academy of Sciences of the United States of America*, 105(42), 16083–8.
- Flege, J. E., Bohn, O.-S., & Jang, S. (1997). Effects of experience on non-native speakers' production and perception of English vowels. *Journal of Phonetics*, 25(4), 437–470.
- Golestani, N., & Pallier, C. (2007). Anatomical correlates of foreign speech sound production. *Cerebral Cortex (New York, N.Y. : 1991)*, 17(4), 929–34.
- Golestani, N., & Zatorre, R. J. (2004). Learning new sounds of speech: reallocation of neural substrates. *NeuroImage*, 21(2), 494–506.
- Golestani, N., & Zatorre, R. J. (2009). Individual differences in the acquisition of second language phonology. *Brain and Language*, 109(2-3), 55–67.

- Hayiou-Thomas, M. E., Harlaar, N., Dale, P. S., & Plomin, R. (2006). Genetic and environmental mediation of the prediction from preschool language and nonverbal ability to 7-year reading. *Journal of Research in Reading*, 29(1), 50–74.
- Hornickel, J., Lin, D., & Kraus, N. (2013). Speech-evoked auditory brainstem responses reflect familial and cognitive influences. *Developmental Science*, 16(1), 101–10.
- Jakoby, H., Goldstein, A., & Faust, M. (2011). Electrophysiological correlates of speech perception mechanisms and individual differences in second language attainment. *Psychophysiology*, 48(11), 1517–31.
- Jin, Y., Díaz, B., Colomer, M., & Sebastián-Gallés, N. (2014). Oscillation encoding of individual differences in speech perception. *PloS One*, 9(7), e100901.
- Johnson, J., & Newport, E. (1989). Critical period effects in second language learning: The influence of maturational state on the acquisition of English as a second language. *Cognitive Psychology*, 21, 60–99.
- Kovas, Y., & Hayiou-Thomas, M. (2005). Genetic Influences in Different Aspects of Language Development: The Etiology of Language Skills in 4.5-Year-Old Twins. *Child Development*, 76(3), 632–551.
- Perani, D., & Abutalebi, J. (2005). The neural basis of first and second language processing. *Current Opinion in Neurobiology*, 15(2), 202–6.
- Sebastián-Gallés, N., & Díaz, B. (2012). First and Second Language Speech Perception: Graded Learning. *Language Learning*, 62, 131–147.
- Weber-Fox, C. M., & Neville, H. J. (1996). Maturational Constraints on Functional Specializations for Language Processing: ERP and Behavioral Evidence in Bilingual Speakers. *Journal of Cognitive Neuroscience*, 8(3), 231–56.
- Young, J. P., Lader, M. H., & Fenton, G. W. (1972). A twin study of the genetic influences on the electroencephalogram. *Journal of Medical Genetics*, 9(1), 13–6.

- Bialystok, E., & Hakuta, K. (1999). Confounded age: Linguistic and cognitive factors in age differences for second language acquisition. (T. & F. E-Library, Ed.) *Second language acquisition and the critical period hypothesis*. Lawrence Erlbaum Associates, Inc.
- Bishop, D. V. M., Adams, C. V., & Norbury, C. F. (2006). Distinct genetic influences on grammar and phonological short-term memory deficits: evidence from 6-year-old twins. *Genes, Brain, and Behavior*, 5(2), 158–69.
- Byrne, B., Olson, R. K., Samuelsson, S., Wadsworth, S., Corley, R., DeFries, J. C., & Willcutt, E. (2013). Genetic and environmental influences on early literacy. *Journal of Research in Reading*, 29(1), 33–49.
- DeKeyser, R. (2000). The robustness of critical period effects in second language acquisition. *Studies in Second Language Acquisition*, 22(4), 499–533.
- Díaz, B., Baus, C., Escera, C., Costa, A., & Sebastián-Gallés, N. (2008). Brain potentials to native phoneme discrimination reveal the origin of individual differences in learning the sounds of a second language. *Proceedings of the National Academy of Sciences of the United States of America*, 105(42), 16083–8.
- Flege, J. E., Bohn, O.-S., & Jang, S. (1997). Effects of experience on non-native speakers' production and perception of English vowels. *Journal of Phonetics*, 25(4), 437–470.
- Golestani, N., & Pallier, C. (2007). Anatomical correlates of foreign speech sound production. *Cerebral Cortex (New York, N.Y. : 1991)*, 17(4), 929–34.
- Golestani, N., & Zatorre, R. J. (2004). Learning new sounds of speech: reallocation of neural substrates. *NeuroImage*, 21(2), 494–506.
- Golestani, N., & Zatorre, R. J. (2009). Individual differences in the acquisition of second language phonology. *Brain and Language*, 109(2-3), 55–67.
- Hayiou-Thomas, M. E., Harlaar, N., Dale, P. S., & Plomin, R. (2006). Genetic and environmental mediation of the prediction from

- preschool language and nonverbal ability to 7-year reading. *Journal of Research in Reading*, 29(1), 50–74.
- Hornickel, J., Lin, D., & Kraus, N. (2013). Speech-evoked auditory brainstem responses reflect familial and cognitive influences. *Developmental Science*, 16(1), 101–10.
- Jakoby, H., Goldstein, A., & Faust, M. (2011). Electrophysiological correlates of speech perception mechanisms and individual differences in second language attainment. *Psychophysiology*, 48(11), 1517–31.
- Jin, Y., Díaz, B., Colomer, M., & Sebastián-Gallés, N. (2014). Oscillation encoding of individual differences in speech perception. *PloS One*, 9(7), e100901.
- Johnson, J., & Newport, E. (1989). Critical period effects in second language learning: The influence of maturational state on the acquisition of English as a second language. *Cognitive Psychology*, 21, 60–99.
- Kovas, Y., & Hayiou-Thomas, M. (2005). Genetic Influences in Different Aspects of Language Development: The Etiology of Language Skills in 4.5-Year-Old Twins. *Child Development*, 76(3), 632–551.
- Perani, D., & Abutalebi, J. (2005). The neural basis of first and second language processing. *Current Opinion in Neurobiology*, 15(2), 202–6.
- Sebastián-Gallés, N., & Díaz, B. (2012). First and Second Language Speech Perception: Graded Learning. *Language Learning*, 62, 131–147.
- Weber-Fox, C. M., & Neville, H. J. (1996). Maturational Constraints on Functional Specializations for Language Processing: ERP and Behavioral Evidence in Bilingual Speakers. *Journal of Cognitive Neuroscience*, 8(3), 231–56.
- Young, J. P., Lader, M. H., & Fenton, G. W. (1972). A twin study of the genetic influences on the electroencephalogram. *Journal of Medical Genetics*, 9(1), 13–6.