UNIVERSITY OF MILANO-BICOCCA

DEPARTMENT OF PSYCHOLOGY

DOCTORAL PROGRAM IN EXPERIMENTAL PSYCHOLOGY,
LINGUISTICS AND COGNITIVE NEUROSCIENCE

CICLE XXVIII

LEXICAL PREDICTION MECHANISMS IN EARLY BILINGUAL SPEAKERS

PhD CANDIDATE:

FRANCESCO GIANNELLI

774636

SUPERVISOR:

PROF. MARIA TERESA GUASTI

CO-SUPERVISOR:

PROF. NICOLA MOLINARO

To Julia

Table of contents

Abstract

Language comprehension is largely supported by predictive mechanisms that account for the ease and speed with which communication unfolds. Both native and proficient non-native speakers can efficiently handle contextual cues to generate reliable linguistic expectations. However, the link between the variability of the linguistic background of the speaker and the hierarchical format of the representations predicted is still not clear. In order to shed light on these matters three ERP experiments have been run.

In the first study, we investigated whether the native experience of typologically highly diverse languages (Spanish and Basque) affects the way early balanced bilingual speakers carry out language predictions. During Spanish sentence comprehension, participants developed predictions of words the form of which (noun ending) could be either diagnostic of grammatical gender values (transparent) or totally ambiguous (opaque). We measured electrophysiological prediction effects time-locked both to the target word and to its determiner, with the former being expected or unexpected. Event-related (P200-N400) and oscillatory activity in the low beta-band (15-17 Hz) frequency channel showed that both Spanish and Basque natives optimally carry out lexical predictions independently of word transparency. Crucially, in contrast to Spanish natives, Basque natives displayed visual word form predictions for transparent words, in consistency with the relevance that noun endings (post-nominal suffixes) play in their native language. We concluded that the native language experience largely shapes prediction mechanisms, so that bilinguals reading in their second language rely on the distributional regularities that are highly relevant in their first language. More importantly, we showed that the individual linguistic experience hierarchically modulates the format of the predicted representation.

The central aim of the second project was to find additional evidence that prediction processing is actually hierarchical, so we created an experiment to test the format of the representation on which prediction is based. Here again, participants were either Spanish or Basque natives. Subjects went through a word matching task and a picture matching task in Spanish where they had to indicate whether the word they read or the image they saw matched with the noun they heard. Importantly, Spanish noun endings could be either

diagnostic of the grammatical gender (transparent), or totally ambiguous (opaque). To detect prediction, we measured the ERPs time-locked to the gender-marked determiner preceding the predicted noun. In the word matching task both groups showed a negative (340 ms) effect on the determiner. Crucially, Basque natives displayed an earlier positivity (150 ms) for determiners preceding transparent nouns. In the picture matching task both groups showed an early activation (160 ms) for determiners preceding opaque nouns. On determiners preceding transparent nouns Spanish natives had an activation at 330 ms, while Basque natives displayed a later one (460 ms). We conclude that native experience models prediction mechanisms since bilinguals rely on the characteristics of the first language also for predicting in the second one. Furthermore, we provided critical evidence that linguistic prediction is hierarchical, with at least two different levels of operation differently represented in terms of neural timing, that are able to communicate at an interface stage.

In the third experiment we analysed the effects that language can have on “low-level” perceptual functions. In particular we explored the possibility that second language prediction influences the perception of the stimuli. From past behavioural and ERP studies emerged that words are more effective at activating concepts compared to non-verbal sounds: people are faster and better in recognizing a picture of a cat after hearing “cat” than after hearing a meow. Two groups of Spanish-Basque and Basque-Spanish speakers with the L2 learnt at three years old took part to the experiment. Subjects listened to a word representing an animal or an object, or they heard the corresponding non- verbal sound. The auditory stimulus was followed by a picture that could match or mismatch with it. Participants had to indicate whether the image matched or not with the stimuli, while ERP signals were measured on the image onset. From the analysis emerged a strong P100 effect on the image coming after the words, but not on those coming after the non-verbal sounds. Crucially this effect was present only in the Spanish-Basque speakers, and not in the Basque-Spanish ones. In the Spanish language, Basque natives were not able to activate visual predictions like Spanish natives did. This outcome is further evidence that the first years of linguistic experience shape the visual prediction abilities. Contrary to the previous studies, here native experience does not determine an advantage for L2 learners, who have been found to be impaired in these fast prediction mechanisms.

Chapter 1

Prediction in Language Comprehension

1. Discovering language prediction

It is widely assumed that anticipatory processing plays a fundamental role in human cognition.

While external information from the world is constantly perceived and processed by the senses, the human brain generates "top-down" predictions based on memory associations created from previous experience to actively and efficiently facilitate the integration of the inputs (Bar, 2007). For instance, by observing and anticipating the trajectory of a ball, a tennis player can adjust his position and successfully hit it with the racket.

Within a variety of neural domains, the understanding of how predictive processing benefits a system is relatively straightforward. With the study of language processing, however, such benefits could not be immediately clear. Traditionally, prediction has been considered an inefficient and cognitively expensive processing mechanism in the domain of language comprehension, where there are many possible ways for communicating a single thought, and the chances of mispredicting are consequently very high.

The idea that language users might predict linguistic features has not been part of the generative grammar tradition. Since the cognitive revolution in the 1950s, linguists and psycholinguists (e.g., Jackendoff, 2002; Morris, 2006) have emphasized the fact that one of the most significant characteristics of human language is that users can understand and produce an infinite number of phrases and sentences they have never experienced before. All possible utterances cannot be stored in the brain: they have to be built on the fly using structural principles. As a consequence, with infinite options available as each new word of an unfolding sentence is encountered, predicting what comes next was not considered a viable strategy, except on the very infrequent cases when contextual constraints were atypically high (Stanovich & West, 1979).

Furthermore, modular views of language processing (e.g., Fodor, 1983; Forster, 1979) have considered language comprehension to be a context invariant, "bottom-up" driven process, with information from one neural domain unlikely to proactively influence processing in another domain before it had been completely processed.

In sum, a combination of these reasons, and the lack of contemporary on-line techniques of investigation has likely contributed to general absence of exploration of anticipatory comprehension in language historically.

Nonetheless, there is now plenty of evidence that language processing is predictive indeed.

Language comprehension involves processing of a noisy sensory input to identify structures (e.g., phonemes, syllables, words, syntax) and integrating them with social contexts, with world knowledge, and with each other, in order to create meaning.

Early language processing models involved some kind of memory buffer by which the elements of the sentence were provisionally stored for later integration (Abney, 1989; Carroll & Bever, 1978; Daneman & Carpenter, 1983; Just & Carpenter, 1980; Kintsch & van Dijk, 1978; Mitchell & Green, 1978; Pritchett, 1991). Since the 1970s, however, the consensus has been that sentence processing is continuous and notions of buffering gradually opened to a more incremental view, where words were integrated sequentially into the sentential context as they were perceived and identified, with gradual accretion of meaning in the mind of the comprehender (e.g., Altmann & Steedman, 1988; Kutas & Hillyard, 1983 ; Marslen-Wilson, 1975; Marslen-Wilson & Tyler, 1980; Pickering, 1994; Steedman, 1989; Traxler, Bybee, & Pickering, 1997; Tyler & Marslen-Wilson, 1977).

More recently, however, there have been many investigations suggesting that language processing may not just be incremental; instead, evidence from various sources has sustained the view that sentential constraint starts to apply its influence even *before* elements have been physically perceived. The debate moved to the predictive role of context[1] in producing expectancies during sentence processing, and a major topic in the sentence processing literature has been whether information about particular words or their features is pre-activated in on-line sentence processing as a product of top-down contextual processing (prediction view), or whether word processing is stimulus driven (bottom-up), triggered by the input, and started only after the physical stimulus has been perceived (integration view). In other words, what some researchers take as evidence for neural pre-

---

[1]*By context we refer to knowledge, stored at multiple pieces within memory about the conceptual features that are necessary (Chomsky, 1965; Katz & Fodor, 1963), as well as those traits that are most likely to be associated with a particular semantic and thematic role of an individual event or state. This knowledge might also include the necessary and likely temporal, spatial, and causal relationships that link multiple events and states together to form sequences of events.*

activation, others take as a sign of the ease or difficulty in integrating words into sentence upon, but not before, their occurrence.

1.1 Integration or Prediction?

A landmark in the debate about integration vs prediction is the N400: a widely studied Event Related Potential (ERP) component (see 1.3 for a closer look at the ERP technique).The N400 is a neural negative response peaking at around 400 ms (200–500 ms) after stimulus onset, and it was first reported by Kutas & Hillyard (1980) who considered it to be an index of semantic violation detection. Further studies showed that the N400 is not only an index of semantic processing for integrating meaning and word knowledge, but also a part of the brain's normal response to potentially meaningful events. In general, it can be said that it (negatively) increases as a function of the fit of a word within the on-going semantic context, therefore it is considered a marker of the difficulty of semantic integration processes (Kutas & Federmeier, 2000; Kutas et al., 2000).

In the ERP literature, the effect of sentence context was usually estimated by manipulating the grade of semantic congruency between the context and its upcoming word (Kutas & Hillyard, 1980), the expectedness (Van Petten & Kutas, 1990), or the sentence constraint (Hoeks et al., 2004; Meyer & Federmeier, 2007). Contextual effects are usually determined by the cloze procedure, where participants were invited to complete a sentence with the first word that comes to their mind. The cloze probability of a word denotes the percentage of participants who finished the sentence with that particular word (Taylor, 1953). A widely replicated outcome is that N400 amplitudes are inversely proportional to the cloze probability. For instance, in the sentence "*He liked lemon and sugar in his tea/coffee*" (Kutas & Hillyard, 1984) the higher the cloze probability of a word (*tea*) in a context, the more reduced is the amplitude of the N400 compared to an unexpected word (*coffee*).

In one of the earlier ERP studies arguing for predictive language comprehension, Federmeierand & Kutas (1999) created highly constraining sentence contexts (e.g., "*They wanted to make the hotel look more like a tropical resort. So, along the driveway, they planted rows of …*"). Sentences were completed with one of three different endings: (1) the expected item, as determined by cloze probability norming (Taylor, 1953; e.g., *palms*), (2) an unexpected item from the expected category (e.g., *pines*), or (3) an unexpected item from a different semantic category (e.g., *tulips*).

Both violation types were unexpected and implausible conclusions for the sentence pairs. Consequently, on an integration account, violations (2) and (3) should be equally difficult to process, as both contain features that are not coherent with the context. For example, unlike *palms*, neither *pines* nor *tulips* is associated with "tropical." Nonetheless, the brain response to these two violations in the N400 region was different. While unexpected items elicited an increased N400 effect relative to the expected exemplars, the N400 response to the two unexpected items differed as a function of their semantic similarity to the expected completions. Violations which shared many semantic features in common with the expected exemplars (2), elicited smaller N400s than did violations which had less semantic feature overlap (3) (Figure 1).

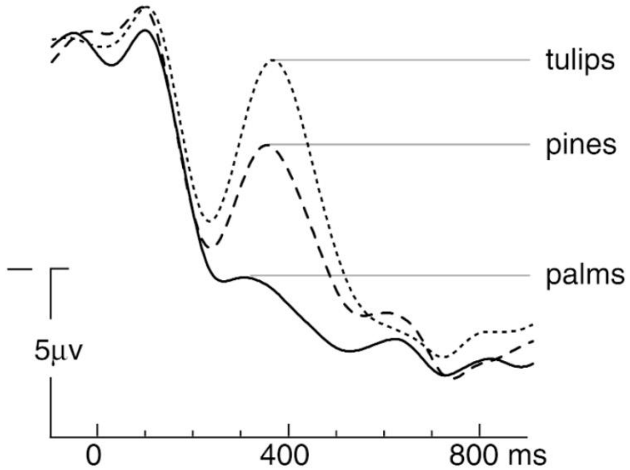


Figure 1:
N400 effect for the three experimental conditions (image from Federmeier & Kutas, 1999).

Authors proposed that this outcome was evidence of predictive processing: violations like *pines* resulted to be easier to process (as indexed by smaller N400 amplitudes) than violations like *tulips*, because of their greater similarity to the predicted, although never actually presented, sentence completion. Still, the observed ERP effects were evident only after target words of interest, leaving them open to the "integration" criticism.

According to these studies, it can be safely stated that language users can use multiple types of information in the context to facilitate the integration of the words when they appear. Nonetheless, facilitation does not necessarily imply predictive pre-activation, because the facilitation of processing words in a sentence can reflect both the ease of integrating the word into context, and the extent to which the context pre-activates specific properties of those words. Therefore, despite the sensitivity of the N400 in response to semantic expectation, it is impossible to define whether N400 modulation during real-time sentence processing means that language users are using context to generate expectancies for upcoming items (prediction view) or whether they are forced by the context itself to devote more or fewer resources to integrating words into sentence representations (integration view).

How can we understand whether language users predictively pre-activate on the basis of their internal representation of context before perceiving bottom-up input at lower representational levels? Can we unmistakably distinguish between integration mechanisms and prediction mechanisms?

Clearly, the argument for language information being predicted would be strongly consolidated if it could be proven that predictions are being formulated before the target words' onset.

1.2 Anticipatory eye movement

Eye-tracking methods have successfully utilized the visual world paradigm to identify preferential looks to visual targets before complete auditory information is available.

As participants listen to a sentence referring to objects while they are presented a visual scene, the eyes move automatically to the objects in the scene as soon as items referring to those objects are heard (Cooper, 1974).

These eye movements are time-locked to the unfolding acoustic signal: when a word is perceived and recognized in the acoustic input the eyes move toward whatever in the visual scene that word could refer to. This paradigm has been successfully used to answer exactly the questions we asked in the previous paragraph.

To understand whether information at a certain point in a sentence could be used to anticipate information at a later point, Altmann & Kamide (1999) showed participants scenes representing, for example, a boy, a cake, and many other inedible objects. Eye movements were recorded as participants heard each sentence. They found that participants looked more toward the cake at when they heard the verb “eat” than when they heard the verb “move”. In the latter case, the verb did not select for one object more than another, and so eye movements were divided between the various objects that were moveable in the scene. Importantly, this increase in eye movements towards the cake when the verb was “eat” occurred before the onset of the post-verbal noun phrase “the cake”. As a consequence, the authors could reasonably conclude that these were anticipatory eye movements. Still, it was not totally clear whether anticipation of the cake at the verb “eat” could reflect what could plausibly be eaten (reflecting only the lexical semantics of the verb), or what could plausibly be eaten by the boy. The experiment could not distinguish if the predictive behavior was connected to the unfolding sentence, or the unfolding word.

In order to disentangle this matter, Kamide et al. (2003), run a slightly different experiment. They contrasted sentences such as “the man will ride the motorcycle” and “the girl will ride the motorcycle” while participants saw scenes representing, for instance, an adult man, a little girl, a motorcycle, and a carousel. The stimuli were designed so that, on the basis of world knowledge and the presented individuals, the man would be more likely to ride the motorcycle than the carousel, and the girl more likely to ride the carousel than the motorcycle. Eye movements were time-locked to assess where the eyes would be directed on hearing “ride”. If the past demonstrations of anticipatory processing had been driven by the verb alone, anticipatory looks should be noticed toward both the motorcycle and the carousel, as both can be ridden. But if they had been driven by the integration of the verb with its linguistic subject, the action with the agent, and the general context, there should be more looks to the motorcycle than to the carousel in the case of “the man will ride…” than in the case of “the girl will ride…,” and vice-versa in the case the girl is the subject. This is indeed exactly what they found.

Based on such research, we can conclude that the language parser is capable of combining visual context with noun/verb semantics to rapidly narrow the possibilities for upcoming input. Processing in sentence comprehension is the result of the integration of each unfolding word with the prior linguistic context, the concurrent visual scene, and general world knowledge. Critically, the process has been proven to happen before the acoustic input is actually perceived: it is the first empirical evidence of language prediction.

1.3 Event Related Potentials in Language Prediction

As we have seen, experimental evidence showing that specific linguistic items or their features are pre-activated before running into the confirmatory bottom-up input is tricky to obtain, due to the difficulty of identifying processing related to an item that has not yet appeared. In addition, language comprehension is characterized by an extraordinary speed and language prediction is (by definition) even faster, therefore anticipation processes are very hard to grasp.

Electrophysiological measures, with a time resolution in the millisecond range, can give us important information in the investigation of the nature and timing of neural networks involved in the anticipation processes, providing physiological correlates of the behavioral results emerged before.

The electroencephalogram (EEG) is a very precise tool that records the voltage variations of electrical activity produced by large populations of brain cells through electrodes placed on the scalp. ERPs are small voltage changes in the electrical brain activity constantly measured by the EEG. They are time-locked to an external event, for example the presentation of a visual or auditory speech stimulus that is supposed to be related to a particular processing mechanism, therefore they are very useful in investigating the perceptual and cognitive processes involved in language comprehension.

Typically, the ERP signal unfolds across positive and negative peaks, called components, associated to the processing of the stimulus. ERP components can vary in latency, amplitude, topographic distribution across the scalp, and polarity. The latency of a deflection (when it reaches its peak amplitude) reflects the time course of the signal from the onset of a specific event. The polarity and the latency at which their amplitude reaches its maximum define the label of the components (N for negative and P for positive components) and reflect the electric field elicited by a dipolar brain configuration. The amplitude has been linked to changes in the strength of the underlying brain response. Topographic distribution is related to direction of the dipole reflected in a specific ERP component. Two components that are similar in polarity and latency but differ in terms of scalp distribution are considered to recruit different brain networks.

ERP components have been found to be sensitive to various kinds of linguistic information; they can reveal and discriminate different sub-processes happening during language comprehension, but how could they be used for detecting language prediction?

DeLong, Urbach, and Kutas (2005) designed a brilliant experiment to test for prediction of a particular feature of the English language. Their study took advantage of the English phonological feature according to which two different indefinite articles, *a* and *an*, are used depending on the initial phoneme of the immediately following word: the article *a* comes before words starting with consonants, while the article *an* precedes words starting with vowel (e.g. "an airplane" and "a kite"). Participants read sentences with a highly constraining context that were continued by a range of more or less expected indefinite article/noun pairings; for example, " *The day was breezy so the boy went outside to fly ...a kite/an airplane in the park* " in which *a kite* is highly expected and *an airplane*, although plausible, is less expected. They recorded ERP effects both on the nouns and on the articles preceding them.

As expected, in line with previous studies there was a significant inverse correlation between N400 amplitudes in the ERPs to the target nouns and their offline cloze probabilities: the higher a word's cloze probability, the smaller its N400 amplitude. The N400 in response to "kite" was smaller than that to "airplane". However, as previously mentioned, this effect on the noun could have been either a consequence of the surprise at finding a word different than the expected one (prediction), or the result of a more difficult integration of the perceived item into the sentence representation (integration).

Crucially, the authors found larger N400-like responses also to indefinite articles that mismatched the expected upcoming noun (e.g. *The day was breezy, so the boy went outside to fly* ***an*** where the word *kite* is the most expected), before nouns were even perceived. In fact, the size of the N400 to the article was graded by the expected noun's cloze probability (Figure 2). Whereas *kite* and *airplane* differ in meaning, *a* and *an* do not, being distinguished only by their phonological forms, therefore articles are equally difficult to integrate into a given sentence representation.

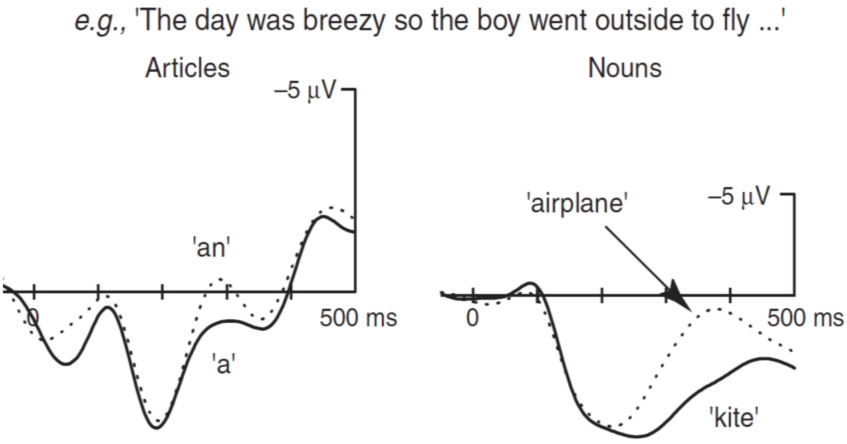


Figure 2:

N400 effects on the article and on the critical noun (image from De Long et al., 2005)

The authors concluded that participants not only anticipated the word that was coming next, but also predicted its phonological features, as reflected in the ERP effect on the article onset. These data are clear evidence of prediction in language at a phonological level, and indicate that people do use sentence context to form graded predictions for specific upcoming words.

What about other lexical features? What happens in languages with a rich morphological system? Will speakers be able to predict other traits of a language?

To answer these questions, Wicha, Moreno & Kutas (2004) took advantage of the obligatory agreement in Spanish between gender-marked articles and nouns to test if individuals make predictions for upcoming words based on sentence context, and if so, whether these contain information about that word's grammatical gender. In their experiment, native Spanish speakers read highly constraining Spanish sentences that contained a gender-marked article followed by a written word. The word target was either the expected continuation of the sentence, or a semantically incongruent continuation of the same gender class as the expected continuation. In half of the sentences, the gender of the article mismatched the gender of the following noun, although participants were not aware about it. The sentences were not grammatically incorrect upon appearance of an article of either gender (expected or unexpected), since the sentences could plausibly have continued with a gender-matching noun even after an unexpected article.

Readers were using sentence context (both meaning and syntactic structure) to anticipate upcoming nouns, including information about their gender, a morphosyntactic property of words in Spanish. The idea was that if any article ERP effect was obtained, it must have reflected the language system's discordance at detecting an article of one gender when it was expecting a noun (and related article) of the other gender. Here follows a set of sample stimuli:

> *Caperucita Roja llevaba la comida para su abuela en …una/un…canasta/corona* .
>
> Little Red Riding Hood carried the food for her grandmother in a basket/crown.
>
> (a)…*una* [feminine] *CANASTA* [feminine] *muy bonita*
> (gender match/semantically congruous)
> (b)…*una* [feminine] *CORONA* [feminine] *muy bonita*
> (gender match/semantically incongruous)
> (c)…*un* [masculine] *CANASTA* [feminine] *muy bonita*
> (gender mismatch/semantically congruous)
> (d)…*un* [masculine] *CORONA* [feminine] *muy bonita*
> (gender mismatch/semantically incongruous)

The pattern of ERPs for target words clearly confirmed the authors' hypothesis. A significant effect of gender expectancy, namely an N400-like negativity, was found at the article preceding the target noun, as well as on the noun. The authors concluded that readers clearly are able to predict not only the semantic and conceptual features of the items that are coming next, but also the morphosyntactic features of these likely forthcoming items. Thus, the language comprehension system is able to make detailed, multifaceted predictions about specific aspects of upcoming words.

Recently Van Berkum et al. (2005) found evidence of anticipatory sentence processing by using characteristic of the Dutch language. In their ERP experiment subjects listened to Dutch sentences that supported the prediction of a specific noun (e.g., *The burglar had no trouble locating the secret family safe. Of course, it was situated behind a* ...). To test whether listeners were indeed anticipating this noun (e.g., *painting*) by the time they had heard the indefinite article, critical stories were continued with a gender-marked adjective whose inflectional suffix did not agree with the noun's syntactic gender. These prediction-inconsistent adjectives elicited a positive deflection in the ERP waveforms, relative to adjectives whose inflection agreed with the lexically stored syntactic gender feature. When this effect was observed the noun had not been presented yet, thus the adjective's inflectional variant was still fully grammatical. This result only makes sense if participants indeed anticipated the discourse-predictable noun.

1.4 Language Prediction mechanisms

The presented studies provide evidence that language processing is predictive: it would be not possible to explain the effects emerged in the previous experiments without assuming that the context influences the state of the language processing system before the bottom-up input is observed.

Of course, there are many obvious potential benefits of prediction in language processing. Prediction could consent faster processing and better efficiency of mental operations. It may solve much of the ambiguity intrinsic in most linguistic utterances and thus reduce memory load.

There are a fair number of psycholinguistic models of sentence processing that include a predictive component. These models are mainly based on ideas proposed by researchers working in a connectionist tradition, and consider prediction to be a graded and probabilistic phenomenon (e.g. Gibson, 2000; Hale, 2006; Levy, 2008; MacDonald, 2013). For the sake of this discussion, here I will not distinguish among these models; instead, I will assume a rather general approach.

A key emphasis of predictive-processing models is an asymmetry between the forward (bottom-up) and backward (top-down) flow of information. Percepts stabilize via a recurrent cascade of top-down predictions reflecting what the system expects given what it already "knows" about the world and about the current context. Predictions are thus combined with incoming sensory data to get at gradually better guesses about the world. But how do speakers create their "internal" knowledge?

The main idea is that during their life, speakers' exposition to their native language determines how they associate words, categories and structures, building an "internal" database of relative frequencies of occurrence and associations, as well as structures (Bar, 2007). These co-occurrences and frequencies are not static but change in time when interacting with other users of the language. While listening or reading, language users construct a syntactic, semantic, phonological and orthographical representation of the input, automatically estimating the several ways in which the current sentence can be continued in a probabilistic fashion. A quite simple example is that the verb *eat* is more likely to occur with an edible object than with an inedible object. This can account for Altmann & Kamide's (1999) findings that listeners anticipatorily move their eyes to a picture of a cake when they hear *The boy will eat the…* . Obviously, prediction becomes hard if more than one, almost equally possible alternative continuations are possible. For instance, if an English speaker hears the fragment *Pick up the can…* in a context in which there is both a candy and a candle, the target object cannot be predicted until more information is available. In this case, participants may either not actively predict, or keep multiple, competing predictions in parallel, which would not guide to an obvious preference for the target object before it appears in the input[2]. If a specific continuation is strongly predicted (meaning that is very likely to occur) but is contradicted by the input or not found in it, the result is a processing difficulty. In this situation, the prediction error, that is the disagreement between the predicted and actual input, is used to modify the options available for upcoming material and to possibly update in memory the frequencies of the structures and associations in question. Language users take advantage of the prediction error to adjust future predictions, with the result of speeding up processing, enhancing communication, adapting to the interlocutor and boosting language learning (Chang, Dell, & Bock, 2006; Federmeier, 2007; Molinaro et al., 2016).

Some authors assign a major role to production mechanisms in predictive language processing. Pickering & Garrod (2013) argue that listeners covertly imitate speakers in

[2]*There remains some debate about whether prediction proceeds in a serial or parallel fashion. According to the first, if the bottom-up input mismatches the predicted structure, the parser reanalyzes and goes to on to the next possibility (e.g. Traxler, Pickering, & Clifton, 1998). In contrast, in parallel models multiple predictions with different probabilistic degree are computed by the parser: if the bottom-up input is inconsistent with the predictions, probabilities are reweighted (Jurafsky, 1996; Levy, 2008, Traxler, 2014). In a slightly different view, multiple predictions are kept active but with different levels of activation (Federmeier & Kutas, 1999; Molinaro et al., 2013).*

order to predict the speaker's upcoming utterances by using production models. These “predictions-by-simulation” are assumed to be impoverished representation rather than fully implemented production representations.

Chang and colleagues (2006) developed a dual path model consisting of two systems. The meaning system is involved in the binding of concepts and event roles; the sequencing system is an error-based learning mechanism. These two systems converge on a word output layer to ensure the right timing of the production of words. Importantly, the constructs required for the production of sentences arise from learners' predictions about upcoming words. Learning therefore occurs when the model's production-based predictions are compared against productions generated by others. Similar to predictive accounts in perception, prediction error is assumed to result in adjustments to the system which generated the predictions.

Mani & Huettig (2012) consider prediction error so fundamental that they suggested that learning to speak in infants arises directly from learning to predict by mispredicting. They presented 2 year old children with audio sentences like *The boy eats a big cake* or *The boy sees a big cake*. While hearing these sentences the toddlers were looking at two-object displays (e.g., a cake and a bird). The subjects predicted and fixated the target object before it was mentioned when they heard the semantically-constraining verb (*eats*) but not when they heard the neutral verb (*sees*). Crucially, children's anticipatory fixations were significantly correlated with their productive vocabulary size: kids with large production vocabularies but not those with small production vocabularies anticipated the target objects. These data suggest that production-based prediction mechanisms may be important in early language development.

To sum up, the “anticipatory processing” approach unifies the idea of parsing, learning, and exposure: during processing, predictions are made as to what is most likely to come in the input. These predictions are based on the frequency with which certain combinations of items have been encountered during previous exposure. If a prediction is not correct, the error is used to adjust the frequencies stored about the words and structures used.

According to anticipatory frameworks, the degree to which a listener or reader can predict upcoming material depends on many interdependent factors, all of which are subject to individual differences due to life experience. If this is the case, what about people who are exposed to more than one language? How would anticipation mechanisms work?

Would second language speakers predict an L2 in the same way monolingual speakers do with the native language? Is prediction even possible in a second language?

2. Prediction in Second Language Processing

Research on second language (L2) anticipation mechanisms represents a fundamental approach for comprehending the general mechanisms underlying language prediction in general.

Non-native speakers differ from native speakers in the nature of the language they are exposed to, and the amount of exposure they receive. The linguistic environment determines the frequency biases that a language user stores and utilizes: the quantity and the quality of the L2 input users are exposed to is much poorer compared to that of a native speaker. It is often the case that second language learners hear the L2 exclusively as produced by other L2 speakers (Selinker, 1972). Also, when L2 users are not immersed in the L2 environment the exposure is limited to textbooks and clearly pronounced and written language, which is very different from the input a native speaker receives.

As a consequence, L2 speakers are different from L1 speakers in the frequency they are exposed to specific combinations of words and structures: this leads to altered frequency biases, henceforth differences in predictive processing. The dissimilarity in frequency has effects both on L2 sentence processing and on the kind of predictions L2 learners may make: a weaker representation of the distributional properties of a language can either trigger weaker predictions, or predictions that are different from those native speakers make (Gollan et al., 2011; Gollan, Montoya, Cera, & Sandoval, 2008).

A second fundamental concept is that L2 speakers have to deal with a dominant, native language in addition to their second language. Several studies on lexical processing have revealed that even when only one language is relevant, bilinguals activate lexical information in both languages. As a consequence, when bilinguals process words in one language, they also activate connected words in their other language (e.g., Dijkstra, Grainger, & van Heuven, 1999; Spivey & Marian, 1999; Van Hell & Dijkstra, 2002). Furthermore, in the auditory domain this co-activation increases: the misperception and the bad representation of phonological categories in the L2 lead to perceive more words in both

L1 and L2 as similar to the input and get activated (e.g.,Weber & Cutler, 2004). Of course, this substantial activation of competing information may damage predictive processing, because anticipation is more difficult if several, almost equally possible continuations are probable. Therefore, L2 learners could potentially activate more information during processing than monolingual native speakers, so they are likely to suffer more from competition: this may obstruct or slow down the creation of predictions.

In other words, non-native speakers differ from natives in terms of quantity and quality of exposure to the L2: this might possibly lead to reduced proficiency, which is likely to cause impairment in L2 processing mechanisms. Nonetheless, it would seem reasonable to assume that when L2 proficiency is high enough, L2 prediction mechanism can be very much like the ones used in the L1, because the efficiency of the prediction mechanism mainly depends on the proficiency in each language. As a consequence, there is no need to assume qualitative differences between native and non-native speakers; instead, the same mechanisms that affect prediction in native speakers can account for the anticipation mechanisms used by language learners (Kaan, 2014).

Several ERP experiments on second language processing corroborate the hypothesis that L2 processing is not different in quality from L1 processing, but it is in fact a function of the proficiency level. Ardalet al. (1990) presented normal and semantically anomalous sentences to fluent bilinguals in their L1 (French) and L2 (English) and to English monolinguals. An N400 was present in all cases, but it was delayed for the bilinguals' L2. Kutas & Kluender (1991) reported similar results, namely a delayed N400 with a reduced amplitude for the less fluent language of bilingual participants. Weber-Fox and Neville (1996) explored the processing of semantic anomalies within sentences in L2 learners who had acquired their L2 (English) at different ages. The N400 was present in monolinguals and all L2 learner groups. However, bilinguals who learned their L2 after the age of 11 and had a lower proficiency level, displayed a delay in the peak latency. Hahne (2001) showed an N400 for semantically anomalous words in Russian late L2 learners of German, however, with a reduced amplitude. In a parallel study with native Japanese late learners of German with a lower proficiency level, Hahne & Friederici (2001) found no significant difference with regard to the N400 between the Japanese L2 speakers of German and native speakers of German.

In summary, the recurrent observation of a similar delay in latency or a reduced amplitude of the N400 found in the participants was directly connected to their L2 proficiency level (and L2 age of acquisition), meaning they had been less exposed to the L2 they were tested in. Nonetheless the results from ERP studies in the lexical–semantic domain point to similarities between L1 and L2 speakers rather than to differences, suggesting that differences between L1 and L2 processing are quantitative and not qualitative in nature.

According to Steinhauer et al. (2009), ERP patterns associated with language processing change over the course of late L2 acquisition. In the earliest stages of learning, ERPs are insensitive to the distinctive properties of the L2. With increasing proficiency, L2 learners' brain activation profiles approaches that of native speakers systematically. At high levels of proficiency, late L2 learners may show ERP responses that are identical to those of native speakers.

In essence, the distinction between L1 and L2 processing consists mainly in a slowdown or decrease in the efficiency of semantic processing mechanisms, depending on the proficiency level of the L2 speaker. Therefore, we could assume that differences in lexical predictive processing between native and non-native speakers can be accounted for by the same dynamics that drive individual differences in native speakers, without assuming that L2 speakers are intrinsically different from native speakers in terms of anticipation mechanisms.

This final statement opens to a central question: what are the differences between prediction in L1 and L2? In the next paragraph we will have a look at the two main experiments that recently investigated L2 prediction processes and addressed the question.

2.1 Characteristics of L2 anticipation processes

Lower L2 proficiency in bilinguals has been shown to delay the N400 effect, suggesting that semantic processing is slower for reading in L2 than in L1 (for reviews see Kotz, 2009; Kutas, Moreno, & Wicha, 2009 Weber-Fox et al., 2003). However, we also showed that recent accounts have suggested that this N400 component reflecting semantic processing is sensitive to lexical prediction (DeLong, Urbach, & Kutas, 2005; Wicha, Moreno, & Kutas, 2004).

As language processing has been found to be slower and more difficult in the second language compared to the first, L2 users may predict upcoming words more slowly than L1 users. Slower sentence comprehension could be disadvantageous for lexical prediction because online integration and comprehension may not work rapidly or well enough for upcoming words to be pre-activated before their occurrence. In other words, L2 users could be able to use sentence constraints to build representation of meaning of the context, but more slowly than L1 comprehenders. If this is the case, a specific lexical item would be predicted, but too late to influence sentence processing before the critical word is encountered. Consequently, L2 comprehenders would not be able to predict upcoming words fast enough.

An alternative possibility would be that L2 readers rely entirely on integration mechanisms during sentence processing. It may be that L2 users build up representations of meaning incrementally, but that such representations affect critical word integration only once the word is read. In this case, items would be integrated based on contextual information stored in working memory, with no active lexical prediction affecting the integration of the critical word and the preceding article.

In 2013, Martin et al. specifically designed an experiment to solve these matters and explore the potential contribution of lexical prediction to semantic processing facilitation during L2 reading, by measuring the N400 elicited by articles preceding expected or unexpected nouns. The goal of the study was to explore whether L2 users predict upcoming words during sentence reading in the same way as L1 speakers, and if reading in L2 involves active lexical prediction of upcoming words like reading in L1. They used a similar ERP paradigm to DeLong et al. (2005) to investigate lexical prediction in late but high proficient Spanish–English bilinguals (L2 comprehenders) who had acquired their L2 after the age of 8, and English monolinguals (L1 comprehenders) reading English sentences. Sentences ended with (a) an expected noun starting with a vowel (e.g., "Since it is raining, it is better to go out with *an umbrella*"); (b) an unexpected noun starting with a consonant (e.g., "Since it is raining, it is better to go out with *a raincoat*"); (c) an expected noun starting with a consonant (e.g., "She has a nice voice and always wanted to be *a singer*"); (d) an unexpected noun starting with a vowel (e.g., "She has a nice voice and always wanted to be *an artist*"). ERPs were time-locked to the article onset and to the noun onset.

In native speakers, unexpected final nouns elicited greater N400 amplitudes than expected final nouns, replicating previous findings (see, for instance, DeLong et al., 2005, 2012; van Berkum et al., 2005; Wicha et al., 2004). In L2 comprehenders, unexpected final nouns also elicited greater N400 amplitudes than expected final nouns (though significantly smaller than in L1 comprehenders). In both groups, semantic processing of nouns was facilitated by the constraining context. This facilitation of integration is due to the fact that the word matched the semantic information activated by the sentence context.

More importantly, in native speakers, unexpected articles elicited greater N400 amplitudes than expected articles, replicating exactly the findings of DeLong et al. (2005). In sentences with a highly constraining context, L1 comprehenders actively predicted the final nouns. In contrast, L2 speakers did not show the same lexical prediction effect: there was no significant N400 effect on the unexpected articles compared to the expected ones.

The authors concluded that second language users did not actively predict the upcoming nouns to the same extent as native speakers did.

This seemed to be a robust result, but did not take into account neither the typological differences between the languages, nor the impact that cross-linguistic similarities may have on anticipation processes. As we saw before, language users do not seem to anticipate only content words, such as nouns, but also their features, for example readers anticipate both the noun and its gender to realize agreement as the sentence unfolds (Wicha et al., 2004). The limitation of Martin et al.'s (2013) study presented above was the linguistic feature used to address anticipation processes. Indeed, they manipulated the determiner alternation guided by the phonological properties of the noun ("a/an"). Crucially, this phonological feature is not present in the participants' L1 (Spanish). Many studies revealed that L2 syntactic processing is more difficult when L1 and L2 syntactic rules differ (Foucart & Frenck-Mestre, 2011; Tokowicz & Mac-Whinney, 2005). Therefore, the lack of anticipation processes could have been due to the complexity of computing the phonological agreement between the noun and the determiner, rather than to an absence of semantic anticipation processes. It might be that participants did predict the noun, but were not successful in applying the phonological agreement rule, and that is why prediction mechanisms were not noticed. Hence, this study left open the possibility that anticipation processes are still possible in L2 processing.

Foucart and colleagues (2014) thought that the best way to solve this matter was to investigate the presence of anticipation processes using a linguistic feature that exists both in L1 and in L2, and follows the same agreement rule. They compared Spanish monolinguals and French-Spanish late bilinguals, gender agreement rules being similar in both Spanish and French. Also, they included a group of Spanish-Catalan early bilinguals who performed the task in their L1 (Spanish). This group was included as a second control group to explore if the bilingual nature of the participants could have an impact on the anticipation processes.

The experiment shadowed Wicha et al. (2004): participants read sentences in Spanish where the context was manipulated in a way that the critical noun and its preceding article could be expected or unexpected. Critically, the gender of the expected noun was the opposite of the unexpected ones, and their corresponding articles agreed with them (e.g., *The pirate had the secret map, but he never found the* [masculine] *treasure* [masculine] / *the* [feminine] *cave* [feminine]).

Interestingly, anticipation effects in late bilinguals emerged from this experiment. ERPs recorded on the unexpected article elicited a more negative N400 effect than the expected articles. In addition, the same N400 effect was found in all three groups of participants, regardless of the L1 they spoke, their L2 proficiency or their age of L2 acquisition. Consistent with this conclusion and with the previous experiments, in all groups there was also an N400 modulation on the unexpected nouns in comparison with expected ones. Thus, the modulation of the ERP components associated with the article and the noun clearly suggests that anticipation processes are at play in the course of sentence comprehension in a second language: bilinguals can predict upcoming words in the same way monolinguals do.

In sum, the difference in L1 and L2 anticipation processes observed in Martin et al. (2013) seems to have been caused by difficult processing in L2 and not by a fundamental difference between L1 and L2 anticipation processes. These results suggest that, like in L1, both integration processes and active prediction mechanisms are at play during L2 sentence reading.

Altogether, the findings of these two studies suggest that language similarity might be the key factor, as well as immersion duration. Nevertheless, given that linguistic processing can become native-like as proficiency increases even when the L1 and the L2 are lexically and syntactically apart (e.g. Dussias, 2003, 2004; Rossi, Gugler, Friederici, & Hahne, 2006), it could be foreseen that highly proficient bilinguals would anticipate upcoming words like monolinguals even when their two languages present cross-linguistic differences.

2.2 The research questions

The two experiments discussed in the previous paragraph show two opposite outcomes.

In the first, L2 comprehenders are not able to predict the features of the L2 words that are going to come (Martin et al., 2013). In the second experiment, L2 speakers predict words' features in the same way native speakers do (Foucart et al., 2014). The main difference between the studies is the language feature analyzed: while in the first experiment the phonotactic agreement feature ("*a*", "*an*") exists in English but not in Spanish, the gender agreement feature between article and noun is present in Spanish, French and Catalan. It is very likely that the reason participants did not show anticipation processing in the first experiment is that the feature they were supposed to predict is not present in their L1. Also, the observations of these two studies could suggest that language similarity might be more important than language proficiency. Indeed, the Spanish-English bilinguals tested in Martin et al. learned their L2 earlier in life were immersed in the L2 country for a longer period than the French-Spanish bilinguals in Foucart et al.

This conclusion leads to fundamental questions: are L2 speakers able to predict the features of the L2 even if those features are not part of their L1? Is prediction fully tuned to the properties of the L2 or can we find traces of influence from the native language? Given a high level of proficiency, is it possible to predict in a native-like way, or do the first years of language experience have an irrecoverable impact on L2 anticipation mechanisms?

In order to answer this questions we designed a set of ERP experiment with bilingual speakers, focused on the link between the language background of a user and the way prediction abilities develop during comprehension. Also, we specifically addressed and controlled the participants' proficiency level and age of L2 acquisition.

Through the analysis of bilingual people, the studies I will present in the next chapters not only aim at finding the differences between L1 and L2 in anticipation processing, but also have the goal of explaining these differences (if any) within a more general and unified theory of the prediction mechanisms underlying language processing.

Chapter 2

The influence of first language acquisition on language prediction in highly proficient bilinguals*

Chapter based on: Molinaro, N., Giannelli, F., Caffarra, S., & Martin C. (under review) Hierarchical levels of representation in language prediction: The influence of first language acquisition in highly proficient bilinguals.

1. Morphological Prediction

In this first study, Basque (L1)-Spanish (L2) and Spanish (L1)-Basque (L2) very early bilinguals read Spanish sentences. We tested participants who were highly proficient in Spanish but were exposed either to Spanish or to Basque before the age of 3. We avoided the comparison between monolinguals and bilinguals since this latter group has a huge amount of competing linguistic information (L2) that can alter the prediction processing dynamics as compared to monolinguals. The present design-comparison between two groups of early balanced bilinguals resolves this confound.

The interaction between Spanish and Basque was considered as highly informative to address our research questions. Spanish and Basque are two rich-morphology languages that are typologically very different on a large number of dimensions (mainly lexical and syntactic). A relevant difference is the way in which these two languages instantiate the relation between content and function words. For example, determiners (articles, quantifiers and prepositions) precede their nouns in Spanish (*la mesa*, "the table"), while in Basque these function words are consistently implemented as post-nominal bound suffixes (*mahai-a*, "the table"; de Rijk & de Coene, 2008; for corpus evidence, Gervain et al., 2013). These function elements are relevant, since they are prominent cues for speech segmentation and signal syntactic boundaries within a sentence. This difference makes Basque speakers focus more on the morphological structure of nouns, more specifically on noun endings (both infants, Molnar et al., 2014, and adults, Gervain et al., 2013).

Based on this typological distinction, the present study took advantage of bilingual sensitivity to the "unsystematic" distributional properties of grammatical gender: a feature that is present in Spanish but not in Basque. In fact, grammatical gender of Spanish nouns is an arbitrary feature (either masculine or feminine) that is uniquely assigned to individual lexical items. This feature is informative of structural relations such as the one between a determiner and its head noun (see Foucart et al., 2014, study described above) and it has been used to study lexical prediction during sentence processing (Wicha et al., 2004). Importantly, noun ending information in Spanish is diagnostic of grammatical gender in only two thirds of cases (*–a* for feminine and *–o* for masculine nouns: cue availability, Harris, 1991). However, since there are plenty of irregularities (~1/3 of the nouns are gender opaque – i.e., *flor*, “flower”, is feminine - or gender irregular - i.e., *mano*, “hand”, is feminine), it has been suggested that proficient Spanish speakers do not rely on formal cues (i.e., the *–a*/*–o* noun ending alternation) to compute agreement dependencies involving grammatical gender, but rely on lexical cues (i.e., the gender value that is lexically associated to each individual noun; Caffarra et al., 2014, 2015; Molinaro et al., 2013a; see Gollan & Frost, 2001, for a dual route proposal).

The present study tested the prediction of gender-transparent (e.g., *mes-a*, “table”, feminine noun) and gender-opaque nouns (e.g., *flor*, “flower”, feminine) by focusing on the processing of the preceding article in Spanish. The preceding article did not differ in the two cases (e.g., *la*, “the”, feminine) so that possible differences between the two conditions were mainly due to the prediction of the following noun.

Based on previous studies, balanced Basque (L1)-Spanish (L2) and Spanish (L1)-Basque (L2) bilinguals should show prediction effects independently from their initial language exposure, since both groups tested are highly proficient in Spanish and, more specifically, in grammatical gender processing in Spanish (see Methods section). This hypothesis is supported by available studies that report similar prediction effects for highly proficient bilinguals and native speakers (Foucart et al., 2014; Hopp, 2013) and proposals stating that prediction in L2 for highly proficient speakers should be similar to L1 (Kaan, 2014).

Crucially, the Basque language does not have grammatical gender and its morphological regularities (specifically, post-nominal suffixes) are highly diagnostic of the underlying linguistic structure (Laka, 1996; Rijk & de Coene, 2008). Spanish grammatical gender (with its large amount of irregularities, ~1/3) provides an interesting test case to evaluate whether the native knowledge of Basque differentially affects the specificity of the predicted representation for transparent and opaque words. It is possible that the Basque natives show more sensitivity to transparent gender cues (noun endings) compared to Spanish natives, since the distribution of nominal terminations is statistically relevant in their native language.

In the present experiment we did not expect prediction differences between transparent and opaque words for Spanish natives. Since they extract grammatical gender information from lexical representations independently from the transparency of the noun, they were expected to develop similar lexical predictions in the two cases. For Basque natives it could be hypothesized that prediction effects are also similar for transparent and opaque words. Overall, the effects could be weaker as compared to Spanish natives since their first language does not have grammatical gender (differently from the study by Foucart et al., 2014, in which prediction effects emerged independently of L2 proficiency, but where L1 and L2 were typologically very similar). Since the level of proficiency in Spanish was controlled across groups, this evidence would support the hypothesis that Basque natives predict in the same way as Spanish natives: they mainly rely on the lexical gender of the target nouns (for morphosyntactic integration effects see Caffarra et al., 2014, 2015). This scenario would support the claim that prediction is independently tuned to the distributional properties of each known language. On the contrary, Basque natives could show differences in the prediction of transparent and opaque words. It is possible that Basque natives rely more on word form properties (noun endings) in predicting gender-transparent words. They would thus predict to a larger extent in the case of transparent words compared to opaque words (but still show prediction effects for both gender categories). This would support the hypothesis that the native language properties interact with prediction in L2 even for highly proficient bilingual speakers.

We measured prediction taking advantage of electrophysiology, since this provides the necessary high-temporal resolution to detect pre-target noun effects time-locked to the preceding determiner. Similar to previous studies (discussed above) we recorded the ERPs time-locked to the target expected determiner as compared to an unexpected determiner (with opposite gender) that introduced a non-anomalous unexpected noun (*En el mapa que tenían los piratas la cruz indicaba donde estaba el tesoro secreto/la perla mágica*, "In the map that the pirates had, the cross indicated where the secret treasure/the magic pearl was."). In line with previous studies, we expected a larger N400 effect for determiners whose gender does not agree with the value of the expected target noun (but see Molinaro et al., 2008, for an alternative functional interpretation of this effect). Also, we explored ERPs on the critical noun (expected vs. unexpected) to explore the consequences of prediction (or absence of prediction, or deceived prediction) on word integration (Molinaro et al., 2010; in L2 literature see Martin et al., 2013).

In addition to previous studies, we also estimated the oscillatory activity time-locked to the target determiner. Differently from the ERPs, the time-frequency estimation provides complementary evidence about neural activity that is not phase-locked to a target event, but presents a relative amount of jittering (variability in its time-course across trials). Even more important, from a theoretical point of view, there is mounting evidence that oscillatory activity in the beta band (13−30 Hz) plays a relevant role in predictive processing. Wang (2010) initially proposed that feedforward visual processing is mediated by feedback-recurrent connection sending top-down information through a beta-band channel. Bastos et al. (2015; Michalareas et al., 2016, for MEG evidence from humans) analyzed the oscillatory dynamics of the primate visual system employing electro-corticography recordings from grids implanted throughout the whole visual system of monkeys. They reported that the beta band activity was associated with feedback influences from higher processing regions to primary visual regions during pre-stimulus visual processing. Similar proposals have been advanced even in the sentence comprehension domain (Lewis & Bastiaansen, 2015; Molinaro et al., 2016). Consequently, we considered it relevant to estimate the beta band components time-locked to the target (unexpected vs. expected) determiner to quantify the strength of the prediction in our experimental design across groups. Evidence of beta-band modulations in our reading design would indicate that more detailed predictions are at work for a specific group/condition (possibly at the visual word form level based on the timing of the effects).

To sum up, in the present sentence reading study we estimated prediction of gender marked nouns whose ending was gender informative (transparent nouns) or not (opaque nouns). Prediction effects were recorded time-locked to the previous gender-marker determiner that could be either gender-consistent or not with the expected noun. Balanced Basque-Spanish bilinguals who were either Basque or Spanish natives took part in the study, our aim being to evaluate how the native language background affects prediction. Similar prediction effects across differently transparent items and groups of balanced bilinguals would provide evidence for the hypothesis that prediction in a second language is just a matter of proficiency; differential prediction effects, depending on transparency, for the two groups of bilinguals would support the hypothesis that prediction is tuned to the distributional regularities of the native language even in fluently proficient second language speakers.

2. Materials and methods

2.1. Participants

Forty-eight early bilingual speakers took part in the experiment. They were divided in two groups. Twenty-four native speakers of Basque (14 females; age range 18−35, mean: 25, SD: 5.10; Age of acquisition of Spanish: 3.75 y, SD: 1.36) formed the first group. They were first exposed to Spanish after the age of 3 and interacted in Basque with both parents. Twenty-four native speakers of Spanish (19 females) formed the second group (age range: 19−41, mean: 24, SD: 4.54; Age of acquisition of Basque: 4.04 y, SD: 1.57); they started to learn Basque after the age of 3 and interacted in Spanish with both parents. Participants received a payment of 10€ per hour for their collaboration. All subjects were right handed and their vision was normal or corrected to normal. All participants signed an informed consent form before taking part in the study that was approved by the BCBL ethics committee.

2.1.1. General proficiency assessment

In order to participate in the experiment, all participants went through a proficiency evaluation in Spanish and Basque (results in Table 1). On a self-rating of their comprehension levels (10-points scale: 0, unintelligible; 10, native-like) they rated themselves very high in both:

For Spanish comprehension there was no difference, while there was a difference for Basque comprehension. We then tested the vocabulary size of our participants in a lexical decision task (no time constraint) in Spanish and Basque (for details of the Spanish version: LexTALE, Izura et al., 2014; Lemhofer & Broersma, 2012). Both groups showed native-range scores for Spanish, and high proficiency scores for Basque (no difference in Spanish level but difference in Basque level, see Table 1). Finally, participants had to name a set of pictures of increasing difficulty in Spanish and Basque. They had native-like scores in Spanish; in Basque they also had very high scores that differ between groups. After the proficiency test, all participants were rated (based on an interview on a 0-to-5 point scale) as fluently proficient in both languages by the experimenters (who were balanced bilinguals). No participant had a score below 4 in either Basque or Spanish, and there was no group difference. Overall, there was no group difference in the general proficiency assessment of Spanish. It should be noted that all the participants also knew English as a third language. This additional language is not relevant in the present design, since participants were largely more proficient in the other two languages. Their proficiency in English did not differ between groups; proficiency was overall rated as good (LexTALE, score 0−40: 23.53, SD: 5.05; picture naming: 46.06, SD: 8.03; final interview: 3.29, SD: 1.31), but still lower than Spanish and Basque.

Measure		Spanish natives (N=24)	Basque natives (N=24)	
Self-evaluation (0-10)	Spanish	9.79 (0.41)	0.50 (0.97)	*n.s.*
	Basque	8.67 (0.96)	9.67 (0.56)	$p < 0.01$
LexTALE	Spanish (0-60)	55.08 (3.21)	54.47 (3.95)	*n.s.*
	Basque (0-50)	38.03 (4.54)	45.00 (3.40)	$p < 0.05$
Picture naming (0-65)	Spanish	64.67 (0.70)	64.17 (1.09)	*n.s.*
	Basque	54.92 (3.27)	64.54 (0.78)	$p < 0.05$
Interview (0-5)	Spanish	4.73 (0.38)	4.70 (0.46)	*n.s.*
	Basque	4.64 (0.60)	4.85 (0.28)	*n.s.*

Table 1: General proficiency assessment of the participants in the two groups

2.1.2. Grammatical gender proficiency assessment

After the EEG experimental session, we further tested the individual proficiency in processing the grammatical gender of the target Spanish nouns employed in the sentence comprehension task. In two complementary tasks we recorded accuracy and response times (reported in Figure 1). In a gender decision task the participants had to identify the gender of the target items as soon and as correctly as possible. Participants were visually presented with isolated words (120 transparent and 120 opaque; same words used in the EEG experiment). Response side was counterbalanced across participants. They had to press left (or right) for feminine words and right (or left) for masculine words. We analysed the data with a two-way ANOVA considering the within factor Transparency (transparent, opaque) and the between factor Group (Spanish and Basque natives). Both accuracy [$F(1,46) = 12.83$, $p < 0.001$, $\eta^2_G = 0.08$] and response times [$F(1,46) = 13.46$, $p < 0.001$, $\eta^2_G = 0.08$] showed a main effect of transparency. In both groups opaque words were more difficult than transparent words, but there was no interaction involving the Group (Figure 1).

Since we studied language prediction focusing on a determiner-noun gender agreement relation, we also evaluated gender agreement in a grammaticality judgement task. Participants were presented with 240 determiner-noun phrases (60 opaque-congruent; 60 opaque-incongruent; 60 transparent-congruent; 60 transparent-incongruent; same words used in the EEG experiment). They had to decide if the determiner-noun phrase was grammatically correct or not by pressing the left or right key of the keyboard (counterbalanced across participants). A two-way ANOVA (Transparency and Group) showed no significant effects in the accuracy. Response times were slower for the opaque items as showed by the main Transparency effect [$F(1,46) = 5.9$, $p < 0.05$, $\eta^2_G = 0.01$]. In addition, a Group effect emerged [$F(1,46) = 4.41$, $p < 0.05$, $\eta^2_G = 0.09$], indicating that Basque natives were slower in their judgements (Figure 3). The two factors did not interact.

Overall, these two tasks indicate that both groups handle grammatical gender similarly, being sensitive to the transparency factor. A main effect of group was observed in the grammaticality task (but not in the accuracy) indicating that Spanish grammar was more complex for Basque natives. However, the lack of interaction with Transparency (that showed reliable effects in all tasks) does not suggest differential processing of the two types of gender at the syntactic level between the two groups.

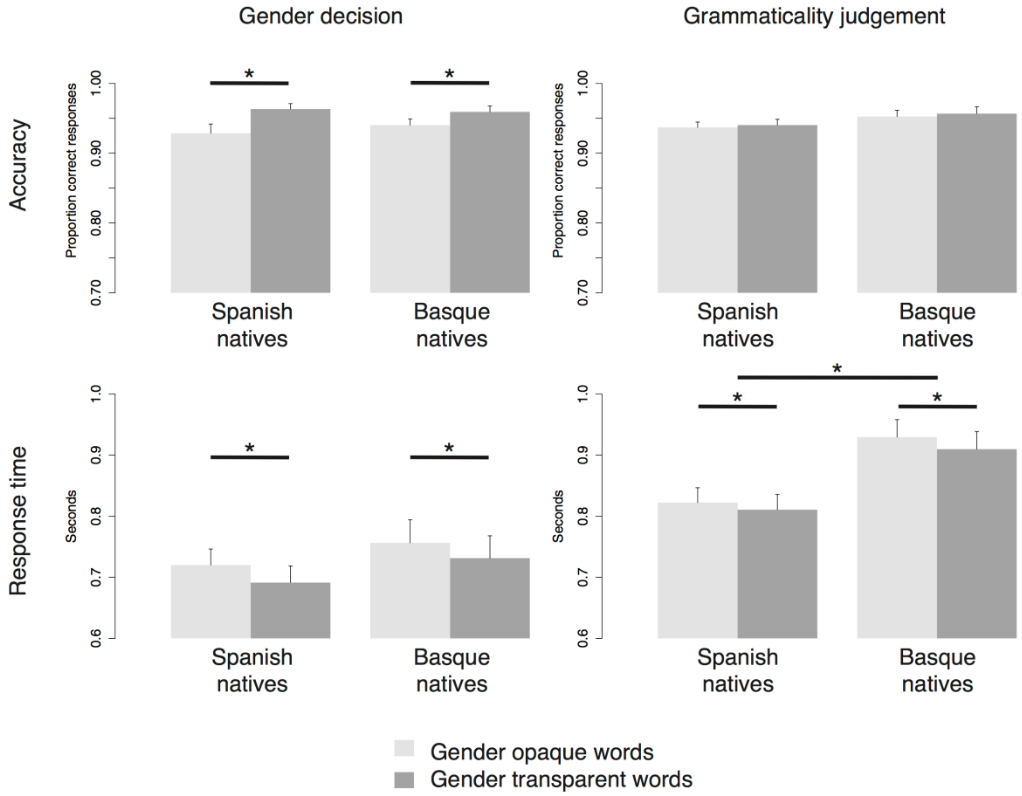


Figure 3: Behavioural results of the grammatical gender proficiency assessment. We here report both the accuracy (proportion of correct responses) and the response times (in ms) for the Gender Decision and the Grammaticality Judgement task in the two experimental groups. Asterisks and horizontal lines indicate the statistically significant differences.

2.2. Materials

Two lists of NPs were created. In the first one, 120 transparent nouns were selected, where 60 were masculine and 60 were feminine, with the masculine nouns ending in "*–o*", and the feminine nouns ending in "*–a*". The second list had 120 opaque nouns that could have different endings. Here as well, half of the nouns were masculine and half were feminine, but the word ending was not informative of the gender value. All nouns referred to inanimate entities. An article ("*el*", "*la*", the; "*un*", "*una*", a) preceded the nouns in both lists. These NPs were employed to construct the sentence stimuli. A hundred and twenty sentence contexts were highly constraining towards a NP (expected condition). The unexpected condition was created substituting the target noun phrase with a different one of opposite gender (but same transparency), resulting in a total of 240 experimental sentences. All the sentences were semantically correct and the target nouns were never in sentence final position. Across sentences the target word (determiner) was on average in position 13.22.

The mean cloze probability of expected and unexpected words was assessed by Basque-Spanish bilinguals (N= 20) who did not take part in the experiment. They had to read sentence contexts and continued them with the very first continuation that came up to their mind. The sentences stopped before the article that should have preceded the word, so that participants were free to use or not the article before the noun. The mean cloze probability for expected nouns and for expected whole NPs was respectively 0.87 (SD: 0.08), and 0.84 (SD: 0.10); the mean cloze probability for unexpected words and unexpected NPs was 0.02 (SD: 0.03), and 0. No cloze-probability differences were observed between sentences preceding opaque and transparent items (all $p > 0.4$).

The 240 sentences were divided in two lists. Each list had 120 sentence contexts followed by 30 transparent expected NPs, 30 transparent unexpected NPs, 30 opaque expected NPs, and 30 opaque unexpected NPs. Each sentence context, as well as each NP, could appear only once in each list in order to avoid repeated presentations, but each sentence context appeared in both lists. We balanced the target words across conditions and lists employing independent ANOVAs (two-way: Expectedness by Transparency). Within each list the target words were balanced (all $p > 0.2$; based on EsPal, Duchon et al., 2013) for grammatical gender, word frequency (log-values: List1: 1.52, SD: 0.56; List 2: 1.52, SD: 0.59), number of letters (List1: 6.09, SD: 1.91; List 2: 5.93, SD: 1.89) and number of neighbors (List1: 6.71, SD: 7.27; List 2: 7.17, SD: 7.55), average position of the determiner (13.22 in both Lists). No differences emerged between lists. Examples of sentences used as experimental material, with the expected vs. the unexpected NP, are reported below:

<u>Transparent item example:</u>

Expected

Acabo de salir de casa y no recuerdo

si | he | cerrado | *la* | *puert-a* | cuando | me | he | ido.

Unexpected

Acabo de salir de casa y no recuerdo

si | he | cerrado | *el* | *armari-o* | cuando | me | he | ido.

[I just left home and I don’t remember if I closed <u>the door</u> (*expected*) / <u>the closet</u> (*unexpected*) when I left]

<u>Opaque item example:</u>

Expected

Prefiero que el te esté muy dulce, | puedes | pasarme | *el* | *azúcar* | por | favor?

Unexpected

Prefiero que el te esté muy dulce, | puedes | pasarme | *la* | *miel* | por | favor?

[I prefer the tea very sweet, could you please pass me <u>the sugar</u> (*expected*) / <u>the honey</u> (*unexpected*)?]

2.3. Procedure

The EEG experiment was run in a soundproof electrically shielded chamber. Participants were seated in a chair, about sixty centimeters in front of a computer screen. Stimuli were delivered with the Presentation software (https://www.neurobs.com/). Participants read sentences displayed in white letters on a grey background. After fixation cross (500 ms), the first part of the sentence was presented as a whole on the screen (average length: 8.2 words, no difference between conditions) for participants to read it. After button press, the second part of the sentence was presented word by word (200 ms +

500 ms inter-stimulus blank interval) until the end. In order to make sure that participants were paying attention to the sentence content, a yes-no comprehension question followed one fourth of the trials: they could answer using the relative buttons on the computer keyboard; response hand was counterbalanced across participants and lists. A brief practice session included three sentences, and the relative yes-no questions. Participants were asked to stay still and to try to reduce blinking and eyes movement to minimum, especially during the word-by-word presentations. Stimuli were presented in three blocks of 40 sentences, with a small break between the blocks. Overall, the experiments lasted one hour and 40 minutes on average.

2.4. EEG recording

Electrophysiological activity was recorded from 27 tin electrodes (Fp1/2, F7/8, F3/4, FC5/6, FC1/2, T7/8, C3/4, CP1/2, CP5/6, P3/4, P7/8, O1/2, F/C/Pz) arranged in an elastic cap (Easycap) according to the extended 10−20 international system. Additional electrodes were placed over the left (on-line reference) and right mastoids. A forehead electrode served as the ground. In addition, four electrodes were placed around the eyes (VEOL, VEOR, HEOL, HEOR) in order to detect blinks and eye movements. Data were amplified (Brain Amp DC) with a bandwidth of 0.01−100 Hz, at a sampling rate of 250 Hz. The impedance of the scalp electrodes was kept below 5 kΩ, while the eye electrodes impedance was below 10 kΩ.

Further data analyses were pursued using Matlab toolboxes (Fieldtrip, Oostenveld et al., 2011; http://www.fieldtriptoolbox.org/) and R (R Core Team, 2015; https://www.r-project.org/). Collected recordings were off-line re-referenced to the average activity of the two mastoids. Raw data were visually inspected and artifacts such as muscular activity and ocular artifacts marked for subsequent rejection. Epochs of interest were computed from −0.3 sec to 2 sec with respect to the determiner onset. Two participants were excluded (and replaced) because more than 20% of the epochs were rejected. On average 5.50 % of epochs were considered artifacts. No difference between conditions and groups emerged in terms of artifact rejection.

2.5. ERP data analysis

After baseline correction (−0.3 to 0 sec) epochs were averaged independently for each condition and subject. We initially focused on a reduced time interval (−0.3 to 1 sec) to select time windows of interest for the analysis of the prediction effects time-locked to the determiner. To this aim we ran for each electrode a point-by-point ANOVA in R considering three factors: Prediction (expected, unexpected), Transparency (transparent, opaque) and Group (Spanish natives, Basque natives). Type-1 error was controlled applying the Guthrie & Buchwald (1991) correction: *p*-values were plotted if they extended consecutively over a period of at least 48 ms (see Janssen et al., 2015). We specifically focused on the main effect and interactions involving the Prediction factor.

Significant interactions were resolved focusing on the average ERP activity across nine groups of electrodes (Left Anterior: F3, F7, FC1; Medial Anterior: Fp1, Fp2, Fz; Right Anterior: F4, F8, FC2; Left Central: T7, FC5, CP5; Medial Central: C3, Cz, C4; Right Central: T8, FC6, CP6; Left Posterior: CP1, P7, O1; Medial Posterior: P3, Pz, P4; Right Posterior: CP2, P8, O2) in the time interval of interest. We ran an ANOVA (Greenhouse-Geisser corrected) with the experimental factors of interest and two additional factors reflecting the electrodes' topographical distribution: Longitude (Anterior, Central, Posterior) and Laterality (Left, Medial, Right). Post-hoc analysis mainly focused on the Prediction effects employing FDR corrected t-tests.

Further analyses were pursued on the longer time-window (−0.3 to 2 sec) to evaluate possible integration effects time-locked to the target noun presentation. Less relevance is given to this latter analysis, since the ERP effects could be affected by earlier modulations time-locked to the determiner. In this latter analysis we focused on the N400 time-interval (five-way ANOVA: Prediction, Transparency, Group, Longitude, Laterality) as an index of successful integration of the target noun in the sentence context (see Martin et al., 2013; Foucart et al., 2014). Proficient readers should definitively show such effects.

2.6. Time-frequency data analysis

The data related to the prediction effects elicited by the determiner were further analyzed focusing on the beta band oscillatory activity (13−30 Hz). Artefact-free EEG data in the time-interval between the determiner and the noun onset (0−0.7 sec) were selected. The time-varying power spectrum of single trials was obtained using a Hanning window approach (400 ms window, 0.5 Hz frequency steps, 5 ms time steps) for the overall frequency range between 2 and 40 Hz (as implemented in Fieldtrip). Power values were expressed as relative change from a baseline interval calculated from −0.3 to −0.05 ms after power estimation single trials were averaged independently for each condition for further statistical analyses and grand-averaged for display purposes.

Statistical significance of the effects was evaluated by means of the cluster-based permutation approach as implemented in Fieldtrip (Maris & Oostenveld, 2007). This approach takes care of the multiple comparison problem by selecting clusters of electrodes, time points and frequency bands that are statistically different between conditions. The initial t-test was set at a probability threshold of 0.05 and the sum of the individual t-statistic in each cluster was employed to determine the cluster statistic. After the randomization procedure (1000 times), clusters exceeding in the highest or lowest 2.5th percentile were considered significant. Pairwise comparisons were focused on the Prediction effect for each transparency level in each experimental group.

3. Results

3.1. Comprehension questions

Participants' responses to the comprehension questions during the EEG session were not significantly different in the two groups ($p > 0.1$). Spanish natives had an average accuracy of 86.98 % (SD: 5.87), while Basque natives' accuracy was 88.28 % (SD: 5.23).

3.2. ERP data

In Figure 2 we report the point-by-point analyses considering the three experimental factors of interest. To better highlight possible prediction effects time-locked to the determiner presentation we report the data in the time window until 1 sec. A strong effect of Prediction emerges in the time interval starting around 250 ms until 400 ms. The effect is evident in most electrodes with a bilateral central-posterior distribution. Interestingly, the effect of Prediction re-emerges similarly across the whole scalp starting around 900 ms. The earlier (250−400 ms) Prediction effect reflects an early N400 effect that is evident across the two Transparency conditions in both levels of the factor Group (Figure 5). The later Prediction effect (> 900 ms) reflects the N400 effect time-locked to the noun following the determiner (Figure 6). In fact, the onset of the following N400 effect is 200 ms after the presentation of the noun. Visual inspection of Figure 4 also reveals a late positive component effect evident after the N400, supporting the claim that the effects observed reflect semantic integration (as in Molinaro et al., 2012).To further validate such analyses and make sure that the Prediction effect is statistically robust for the two groups, we ran the statistics (four-way ANOVA: Prediction, Transparency, Longitude, Laterality) independently for the two experimental groups. Both in the 250−400 ms post-determiner-onset time window [Spanish natives: F (1,23) = 11.81, p< 0.001, η^2_G = 0.05; Basque natives: F (1,23) = 15.31, p< 0.001, η^2_G = 0.03] and in the 200−500 ms post noun-onset time window [Spanish natives: F (1,23) = 32.60, p< 0.001, η^2_G = 0.13; Basque natives: F (1,23) = 9.24, p< 0.001, η^2_G = 0.04] reliable effects of Prediction emerged. When the Group factor is also included in these analyses, no interaction of the experimental factor with Group emerged. In Figure 1 a triple interaction is also visible in the time interval between 170 and 250 ms in central and parietal electrodes. We further explored this effect statistically (five-way ANOVA: Group, Prediction, Transparency, Longitude, Laterality) in this time interval reporting the triple interaction between Group, Prediction and Transparency [F (1,46) = 6.42, p< 0.01, η^2_G = 0.01]. To evaluate such interaction, we ran separate analyses in the two groups. Spanish natives did not show any main effect or interaction with the Prediction factor in this time interval (all *ps*> 0.1). Basque natives, on the other hand, showed an interaction between Prediction and Transparency [F (1,23) = 12.89, p< 0.001, η^2_G = 0.02]. When considering the Prediction factor independently from the two levels of the Transparency factor, the effects were seen to be highly robust for transparent items [F (1,23) = 8.14, $p < 0.01$, $\eta^2_G = 0.02$] but not for opaque items [F (1,23) = 2.38, $p > 0.1$]. This effect is evident in Figure 3 as a decreased positive peak around 200 ms (P200) for unexpected vs. expected transparent items read by Basque natives.

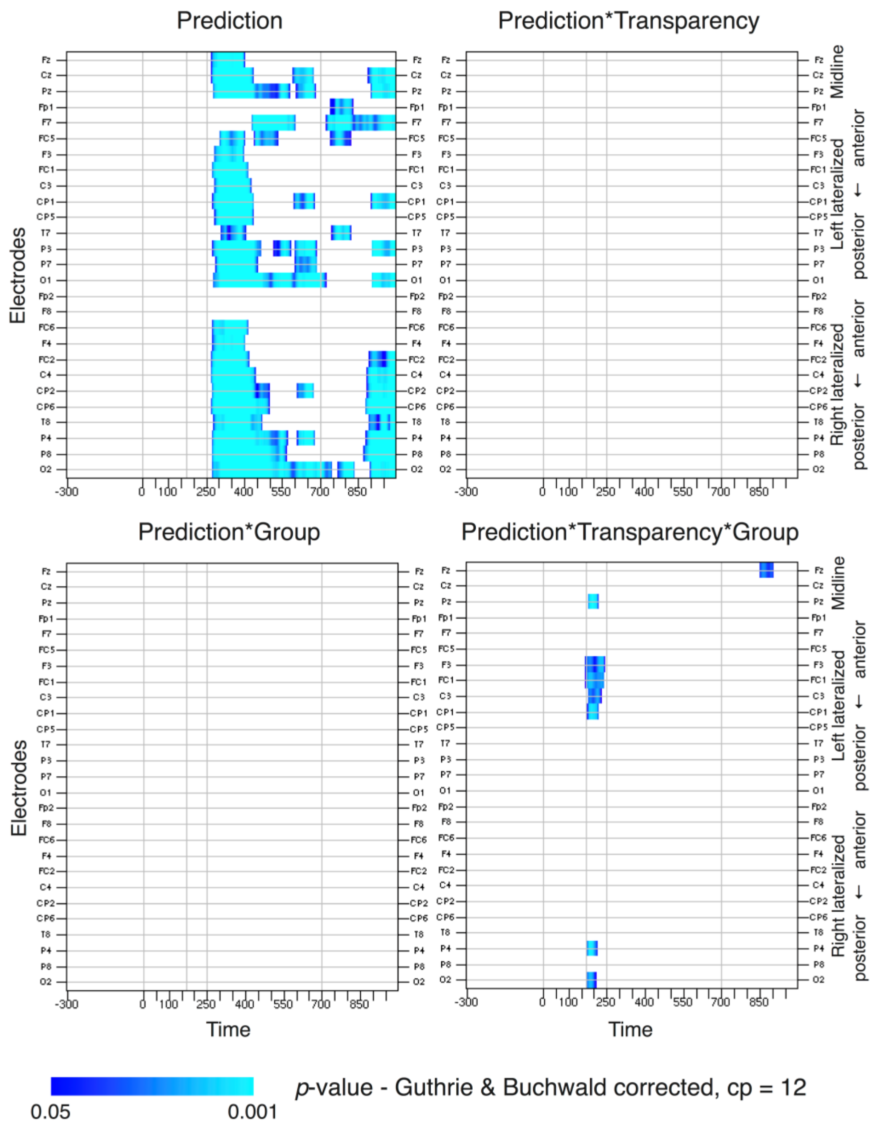


Figure 4: Point-by-point split-plot analysis of variance (Guthrie and Buchwald, 1991, corrected) for each electrode considering the three experimental factors (within: Prediction and Transparency; between: Group). We report the main effect and the interactions involving the Prediction factor. Vertical grey lines at 0 and 700 ms indicate respectively the onset of the determiner and the onset of the predicted noun. Vertical grey lines at 170 and 250 ms indicate the time interval in which the triple interaction emerged.

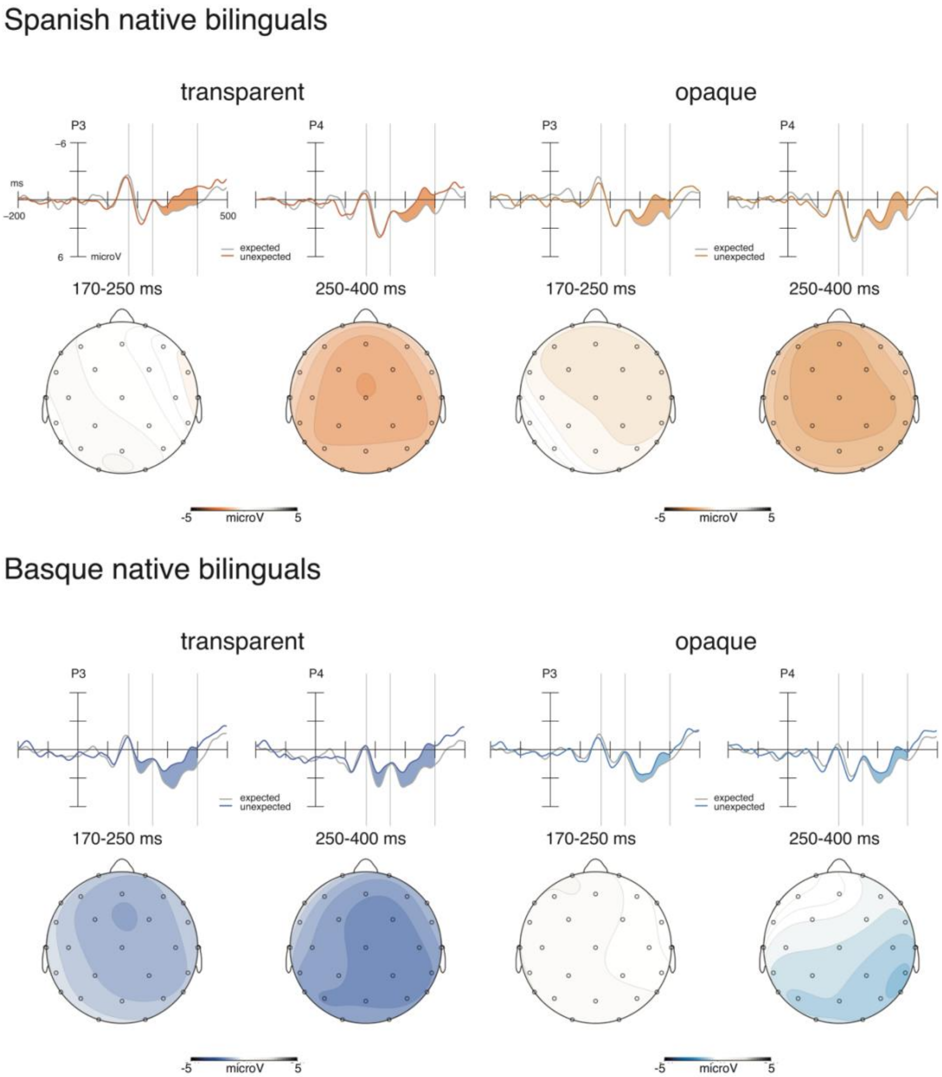


Figure 5: ERPs for the expected and unexpected condition (Prediction effect) time-locked to the presentation of the determiner preceding the predicted noun (plotted in −200 to 500 ms for better displaying the ERP effects). We plotted separately the conditions based on the Transparency of the predicted noun and the Group of native speakers. For each plot we report in the upper panels the waveforms in two representative parietal electrodes (P3, P4; negative values plotted up) where we highlight the time intervals of interest (P200: 170−250 ms; N400 250−400 ms). Shadowed differences are the statistically significant ones. In the lower panels we present the topographical distribution of the difference effect (unexpected minus expected) in the two time intervals of interest.

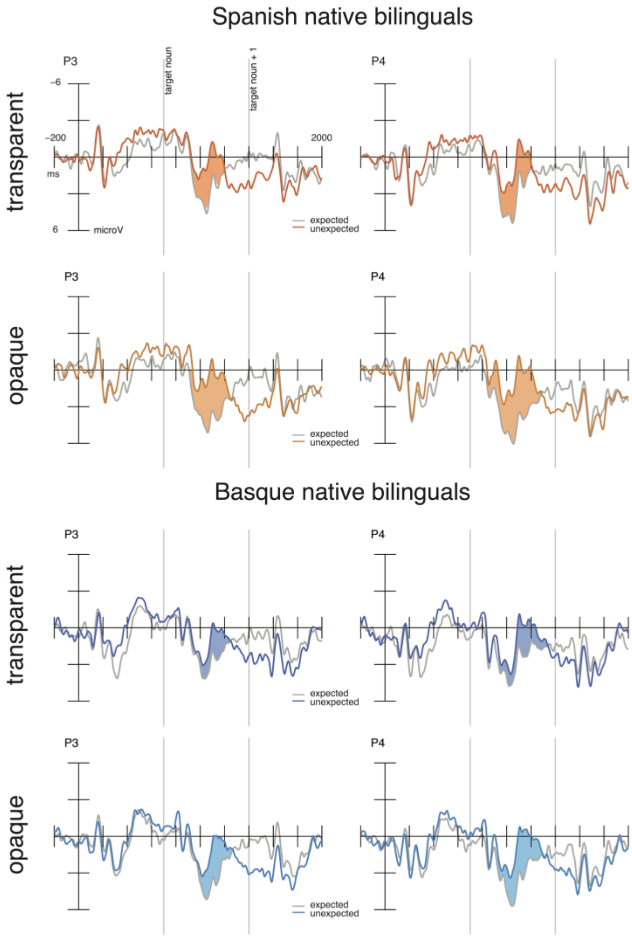


Figure 6: ERPs for the expected and unexpected condition considering the longer time interval (−300, 2000 ms) including both the determiner and the predicted noun in two representative parietal electrodes (P3, P4; negative values plotted up). Shadowed differences reflect the N400 effect post-noun onset, showing successful integration of the prediction. This effect is significant in all conditions.

3.3. Time-frequency data

The oscillatory power analyses pursued in the determiner time-window (0−700 ms) were aimed at evaluating possible beta-band (13−30 Hz) effects as an index of word-form level prediction (Bastos et al., 2015; Michalareas et al., 2016). Across the four unexpected vs. expected comparisons reported in Figure 5 only the one involving the Basque group for transparent items revealed significant results. More specifically, two clusters emerged in this comparison. An earlier one (Cluster 1) mainly involved central electrodes and was significant in the 196−256 ms time interval and between 15 and 17 Hz (lower beta band). The later one (Cluster 2) emerged in the same lower beta frequency band and involved slightly more right-lateralized electrodes between 438 and 496 ms. Both clusters show more power in the low beta band for the unexpected condition as compared to the expected. We further explored other frequency bands (from 2 to 40 Hz and from 20 to 100 Hz by means of a multi-taper approach) but no reliable effects were observed. It is interesting to note that Cluster 1 highly overlapped in time (ms) and space (scalp topography) with the early decreased P200 for this same comparison in the ERP analyses. We further explored the nature of the early oscillatory effect (Cluster 1) by means of Pearson correlation (R software) between the beta band effect and the tasks involving grammatical gender for the Basque natives group. We thus computed for each Basque native participant the transparency advantage, i.e., the difference in response time for opaque minus transparent items in the gender decision task. Then we extracted the difference between the beta power for the unexpected minus the expected condition for transparent items in the most representative electrode (C4 reported in Figure 5). The two measures positively correlated (r (23) = 0.53, $p < 0.001$) showing that the participants who had a stronger sensitivity to transparency in the gender decision task were the ones showing a larger effect in the lower beta range (no correlation for Spanish natives). No relevant correlation was observed involving response times for the grammaticality judgement task.

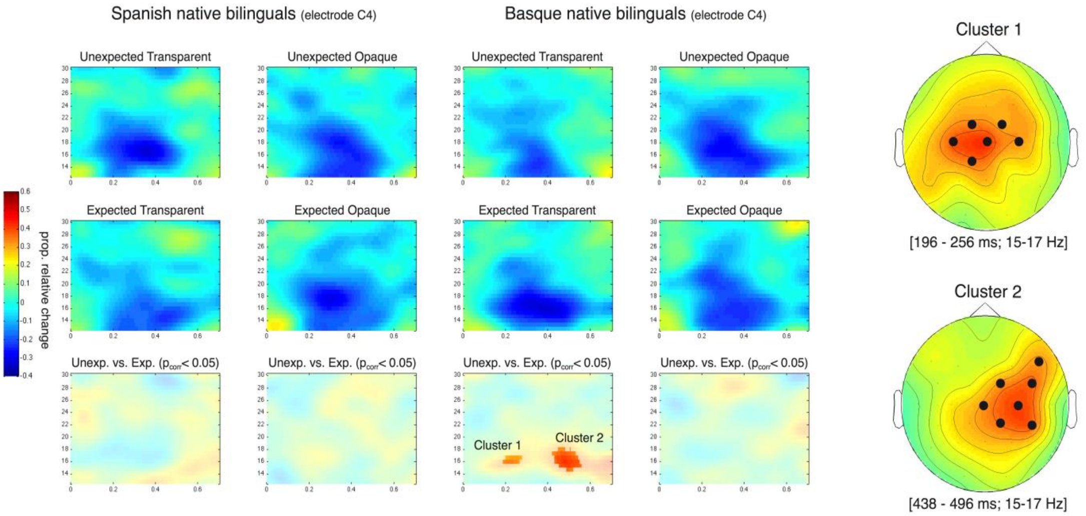


Figure 7: Oscillatory power beta-band activity (13−30 Hz, relative change with respect to the baseline −300, −50 ms) time locked to the determiner preceding the predicted noun (at 700 ms). We report both the beta power decrease for all the conditions and the difference between unexpected minus expected determiner in a representative electrode (C4). Statistically significant time-frequency clusters are not shadowed, showing two significant effects in the 15−17 Hz frequency band for the transparent contrast in Basque natives. The topographical distribution of these two significant clusters is reported on the right, in which the electrodes contributing to the cluster are marked.

4. Discussion

In the present experiment participants read sentences that were highly constraining towards a specific lexical item. Grammatical gender features are encoded in the lexical representation of nouns (Harris, 1991), since each noun has its own grammatical gender (*mes-a*, table, is only feminine; see also Levelt et al., 1999). This makes such a feature a relevant constraint for lexical prediction.

In the present design, when bottom-up information provided by the determiner interacts with top-down information provided by the predicted noun, the two sources of information can either match or not. The timing of the ERP Prediction effect (unexpected vs. expected determiner) reveals the processing stage at which the two representations interact. The unexpected determiner provides a grammatical gender value that contrasts with the information encoded in the predicted lexical item. This "representational contrast" triggers a conflict effect that takes place at the lexical level of processing for Spanish natives, independently of the transparency of the predicted lexical element. The 250−400 ms effect (Figure 3) likely represents a lexical-related negativity (earlier N400) already reported for function words (King & Kutas, 1998; Molinaro et al., 2008; Osterhout et al., 1997). This effect replicates what Foucart et al. (2014) reported for Spanish monolinguals and Spanish-Catalan early bilinguals employing a similar design.

4.1. Early prediction effect for Basque natives

The same lexical effect was observed also for Basque natives for both transparent and opaque items, but it was preceded by an even earlier effect for transparent words. The timing (170−250 ms) of this earlier effect is indicative that the gender value of the determiner mismatches with the gender value of the predicted transparent items at a pre-lexical level. The topographical distribution of this effect indicates that this is a P200 effect, an ERP component classically considered as reflecting visual sensory processes under attentional control (Luck, 2005; see also Molinaro et al., 2013b) showing decreased positivity for items that require increasing attention. The oscillatory evidence (involving the low beta-band channel, as in Bastos et al., 2015; Michalareas et al., 2016; see also Molinaro et al., 2016) further supports the idea that for this last condition we observed a prediction effect involving visual word form representations. In our view, the increased beta power for the unexpected determiner likely reflects an on-line update of the predicted representation: since the determiner is not gender-consistent with the predicted noun, the system inhibits such initial prediction and activates other possible (less predicted) lexical candidates.

The present data thus reveal that Basque natives activate word form-level representations when predicting items whose ending is gender informative. The more sensitive these speakers were to gender transparency of the target items (in the gender decision task, Figure 1) the stronger the word form prediction was (as evidenced by the beta effect). In our opinion, this word-form prediction is not due to a reduced proficiency for Basque natives in Spanish since in the general proficiency assessment the two groups of bilinguals were perfectly balanced for Spanish. In the grammatical gender proficiency assessment we also did not find reliable differences between groups. The slower response times for Basque natives in the grammaticality judgment task would have predicted a later (and not earlier) electrophysiological reaction for this group compared to Spanish natives; importantly, the lack of interaction between transparency and group in the grammaticality judgment task (Figure 1) does not parallel the strong interaction effect we observed in the EEG experiment involving transparency and group (Figure 2). Furthermore, we cannot interpret the similar lexical effect for transparent and opaque items (the one emerging at 250−400 ms) as an index of lower proficiency for Spanish native bilinguals, since this overlaps (in time and scalp topography) with what has been reported for Spanish monolinguals with transparent items (Foucart et al., 2014). Previous studies have already suggested that Spanish speakers do not handle transparent and opaque items differently during sentence processing, given the large amount of irregularities (~1/3) in noun-endings within the Spanish lexicon (Caffarra et al., 2014, 2015; Molinaro et al., 2013a).

We attribute the early effect for Basque natives to their first language acquisition experience. Basque natives in their L1 strongly rely on post-nominal suffixes as highly frequent cues that are employed to bootstrap syntactic representations of Basque during language acquisition (Molnar et al., 2014) and for speech segmentation (as suggested by Gervain et al., 2013, see Gervain & Mehler, 2010). We have pointed out in the introduction that this morphological information stimulates more attention to the word-form properties (and more specifically to the noun endings) in the native language. It is likely that in our experiment, Basque natives activate word-form representation also in Spanish: when the gender value is available at the word form level (for transparent items) an early word form electrophysiological effect emerges; when the gender value could be extracted only at the lexical level of processing (for opaque items) a lexical effect emerges.

4.2. The role of native exposure

The present findings thus speak for the idea that language prediction is mainly tuned to the native language characteristics (*native exposure*). Such properties largely vary across languages and speakers adapt their predictions mainly to the regularities of their native language. As mentioned in the first chapter, recent proposals (Chang et al., 2006; Mani & Huettig, 2012) suggest that prediction is a relevant mechanism through which infants bootstrap the statistical regularities of the language to which they are initially exposed. Based on efficient prediction mechanisms, children can start to develop appropriate language production skills ("Prediction is Production" in the P-chain by Dell & Chang, 2014; see also Pickering & Garrod, 2013; Molinaro et al., 2016). Along the same lines, associative learning theories state that prediction stimulates learning (Rescorla & Wagner, 1972; Schultz et al., 1997) and this is possibly true also for language (Kuperberg & Jaeger, 2016).

This prediction mechanism applies even to an L2 (and for a feature not present in L1) when speakers can properly master it. Here we showed that even when our Basque natives are daily exposed to the statistical regularities (and irregularities as in the case of grammatical gender) of Spanish, they still show signs of the influence of their native language during prediction. It is plausible that the different prediction effects we observed in our two groups (Spanish and Basque natives) are mainly due to the fact that those speakers were exclusively exposed to their native language until the age of three. Consequently, they efficiently learned to predict based mainly on the regularities of their native language, adjusting their prediction mechanisms either to Spanish or to Basque. Predictors initially tuned to Spanish handle transparent and opaque items similarly during prediction, given the reliability of lexical-level grammatical gender processing (Caffarra et al., 2014, 2015), while predictors initially tuned to Basque show more word-form effects even in an L2 when word-form (noun endings) cues are available.

As an alternative explanation for the findings of the present study, it could be argued that while our two groups were proficiency-balanced in Spanish, they were not so in Basque. Even if they both show high proficiency in this language, the Spanish natives were statistically less proficient in Basque than the Basque natives. This could have determined the stronger influence of Basque (L1) on Spanish (L2) for Basque natives, compared to the reduced influence of a weaker Basque (L2) on Spanish (L1) for Spanish natives (*L2 attrition*). We cannot completely exclude this hypothesis in the present study. Nonetheless, to examine this more closely, we selected the five Spanish native speakers that were most proficient in Basque (LexTALE mean score: 44.60; picture naming: 60.20; similar to Basque natives, see Table 1). We averaged their ERPs for the transparent conditions and we did not find quantitative differences in the 170−250 ms interval (amplitude across all electrodes: expected condition: 0.86 microV; unexpected: 0.88 microV; midline electrodes: expected: 1.77 microV; unexpected: 1.73 microV; electrodes showing the triple interaction in Figure 2: expected: 1.51 microV; unexpected: 1.83 microV). This last qualitative analysis suggests that even for the Spanish natives who have higher proficiency scores in Basque there is no P200 expectation effect for transparent words. This does not support the idea that in the present experiment we observed an L2 attrition effect.

4.3. Developing prediction processes

In terms of proposals about L2 prediction, this study does not support hypotheses indicating that prediction in L2 is only a matter of proficiency (i.e., the more you know a language the more native-like your predictions will be; see review by Kaan, 2014). On the contrary, the present study shows that the native prediction mechanisms adapt to the properties of a second language, identifying regularities similar to those available in the native language. This view thus supports the idea that there are no separate domains of prediction, but that it is a unified mechanism looking for useful cues independently of the language processed. This can possibly provide clues and new directions for language learning research. Similar distributional regularities between a native and a new-learned language could serve to boost prediction in a second language and, consequently, facilitate its learning (Kuperberg & Jaeger, 2016; Molinaro et al., 2016).

We would like to emphasise that we do not think that prediction can only develop until the age of 3 and that it cannot further change through experience. Importantly, predictions can flexibly and rapidly adapt to the conditions of a new context (Bar, 2007; Sohoglu & Davis, 2016) by picking up all the available cues to construct an internal representation of the new environment (and develop predictions). However, early experience biases the way in which different cues are weighted to pursue optimal prediction mechanisms in the new experience settings.

4.4. Conclusion

The present study provides new insights into the mechanisms of prediction in sentence comprehension. Taking advantage of the typological distance between Spanish and Basque in early balanced bilinguals, we add an important piece of evidence to the puzzle on how multilingual experience shapes language prediction. Both evoked (P200-N400 prediction effects) and oscillatory electrophysiological evidence (15−17 Hz beta band activity) indicate that prediction can top-down reach the word-form hierarchical level of representation even in a second language. Based on this, we advance the hypothesis that prediction mechanisms are highly influenced by the properties of early language experience.

Chapter 3

Language prediction hierarchies and native language experience*

1. Hierarchical levels of representation in bilingual prediction

The central aim of this second research project was to find additional evidence that prediction processing is actually hierarchical at an activation-timing level, by creating an experiment that could test the format of the representation on which prediction is based. In order to describe the reasoning behind this new study we need to sum up first the outcome of the previous experiment.

As we saw, the Basque language does not have grammatical gender, and its morphological regularities are based on post-nominal suffixes (Laka, 1996; Rijk, 2008). On the contrary, in Spanish, grammatical gender is a feature that is assigned to any inanimate noun, but the noun ending is not always diagnostic of the gender as there is a substantial amount of irregularity. Two thirds of Spanish nouns are gender-transparent (ending in –a for feminine and in –o for masculine nouns), but the remaining nouns are gender-opaque (ending in a consonant, or in a different non-diagnostic vowel) (R.A.E., 2010).

It has been hypothesized that grammatical gender is extracted on the basis of two different routes: one based on the lexical properties (abstract features) of the noun stored in the mental lexicon, and one deriving the gender features from the formal properties of the noun. When the lexical route is not available, language users rely on the formal route to extract gender (Gollan & Frost, 2001).

According to this theory, as Spanish opaque nouns do not have formal cues for gender retrieval, grammatical gender necessarily has to be derived from the lexical route. In contrast, transparent nouns present reliable formal cues (e.g. gender-related noun endings), therefore both the lexical route and the form-based route are utilizable for gender extraction (see Caffarra et al., 2014, for supporting data).

In the previous experiment not only did expected and unexpected nouns differ in gender, but transparency was also taken into account. Participants read sentences that could end with a transparent or an opaque noun, such as *Acabo de salir de la casa y no recuerdo si he cerrado la puerta/el armario"* ("I just got out of the house , and I don't remember whether I closed the [fem] door [fem, transp] / the [masc] wardrobe [masc, transp]") or *Prefiero que el te esté muy dulce, puedes pasarme el azúcar / la miel por favor?* ("I prefer my tea very sweet, would you pass me the [masc] sugar [masc, opaque] / the [fem] honey [fem, opaque] ?").

ERPs recorded on the unexpected determiner (in comparison with the expected) revealed an N400 effect in both groups, independently of the transparency of the predicted lexical element. This result not only replicated Foucart et al. (2014), and demonstrated that highly proficient bilinguals can predict the way monolinguals do, but did so by using a feature, namely gender, which is not available in the L2 of the participants.

In addition, the ERP analysis presented an even more striking outcome. On the unexpected determiners that preceded transparent words, Basque natives showed a P200 effect (starting ~100 ms before the observed N400s), a component classically thought to reflect visual-attention processes (Luck & Hillyard, 1994; Liu et al., 2013; Molinaro et al., 2013; Su et al., 2016).

In order to explain these results we formulated a multifaceted hypothesis. In our opinion, Spanish natives do not rely on formal cues (i.e. *a/o* noun ending) to compute agreement dependencies involving grammatical gender (since ~1/3 of the nouns are gender opaque) (Caffarra& Barber, 2015), but rely on the lexical information stored in the mental lexicon to predict the gender of the word that is coming next, without taking into account the noun transparency, as reflected by the N400 effect.

On the other hand, Basque natives rely more on sublexical, word-form related analysis for the gender prediction of transparent words, as reflected in the P200. They perform the same kind of prediction for opaque words and necessary fail because they have no formal cues to base their prediction on, but they can still switch to the lexical route in order to predict the gender. The reason for this processing difference would reside in the fact that in Basque, Basque speakers are driven by default post-nominal suffix analysis to bootstrap syntactic cues in their mother tongue (Molnar et al., 2014), and they apply the same strategy in the L2.

This hypothesis implies two fundamental issues. First, it provides a step further in the research on multilingual prediction that goes beyond recent paradigms that have worked on the presence or absence of L2 prediction depending on the speakers' proficiency. These results show that the native language experience strongly influences the way prediction is carried out, in other words, the environmental language regularities available during early childhood shape reliance on different levels of linguistic representations for prediction.

This leads to the second point: the experiment provides electrophysiological evidence that language prediction is not a “representation encapsulated” phenomenon (dealing only with high-level linguistic representations) but it flexibly takes advantage of the linguistic representations made available by different cognitive processes across the form-to-meaning (perception-to-abstraction) hierarchy. Predictive coding approaches (Bastos et al., 2012; Friston, 2005; Rao & Ballard, 1999) assume that human interaction with the environment is largely based on internal knowledge-based expectations (*priors*) that, through a top-down process, hierarchically percolate down from memory-based abstract internal representations to sensory regions, thus shaping human perception. So far, experimental evidence of such prediction hierarchy in the language domain has been missing and our recent data provides initial evidence in this direction.

2. The present experiment

The aim of the present study is to test the above hypothesis concerning hierarchical predictions during language processing.

We analysed two similar groups of highly proficient bilinguals: participants were Spanish (L1) - Basque (L2) speakers and Basque (L1) - Spanish (L2) speakers, with the L2 acquired at the age of three years old.

In contrast with the previous experiments, we decided not to use sentences providing a constraining context to trigger lexical prediction. In a sentence context, the effects recorded on the article do not ensure that participants are actually predicting the following content word, but could reflect prediction of the determiner (see Luke & Christianson, 2016), as it plays a relevant sentence-level syntactic role. We thus utilized a word matching task (referred to as WMT in the following) and a picture matching task (PMT), and combined them with the EEG to have the necessary high-temporal resolution to detect anticipation effects.

In two different blocks of the same experimental session, participants read a noun in Spanish on a screen, or they saw the picture of a noun on the screen (e.g. *cuchillo* (knife) or the image of a knife). These stimuli (referred to as *predictors*) were followed by a voice saying a noun phrase (NP, a noun – the *predictee* – preceded by a determiner). Here the noun could be congruent or incongruent with the word that they read or the image they saw. The gender of the incongruent noun was the opposite of the gender of the presented stimuli;

furthermore, nouns could be transparent or opaque. The determiner preceding the noun in the NP could be congruent, incongruent or neuter in relation to the stimuli, but it was always in agreement with its noun. Importantly, the determiner could also be gender-congruent with the predictee but it preceded a mismatching noun in some trials.

Participants had to indicate whether what they read or what they saw (the predictor) corresponded to what they heard (the predictee).

We recorded the ERPs time-locked to both the determiner and the noun that the participant heard. The reasoning behind the experiment comes from straightforward psycholinguistics models. In language processing, when users read a word, sensory information feedforwards activation to the sub-lexical (abstract orthographic) word-form features of that item (reliably after 200 ms post-stimulus onset), and then to the lexical/semantic information related to that word (after 350 ms, Grainger & Holcomb, 2009; for theoretical proposals see Grainger & Jacobs, 1996; McClelland & Rumelhart, 1981). In contrast, after picture presentation, the visual information directly activates the lexical/semantic information related to the image (after 200 ms post-stimulus), from which inflected word-forms can be derived (after 300 ms), similarly to what happens during language production (Indefrey & Levelt, 2004; Strijkers & Costa, 2016; for a parallel between prediction and production see Dell & Chang, 2014; Pickering & Garrod, 2013; Molinaro et al., 2016)**.**

The rationale of the present study is to track the time course of the mismatching effect (if we observe any) time-locked to the determiner, depending on the predictor format (either a written word or a picture). In order to perform the matching task, participants pre-activate the word they expect to hear (the predictee), according to the item they perceived (the predictor). The gender values of the predictor will be consequently activated in time depending on the hierarchical level in which they are represented. In the case of WMT, we expect that the predictor (the written word) will sequentially activate the following representational levels, i.e., visual > sub-lexical > lexical. Importantly, access to lexical representation is mediated by sub-lexical processing. If the gender value is represented at the sub-lexical level, the mismatching gender information provided by a mismatching determiner will trigger an update of the prediction earlier in time compared to a lexical-level gender representation. The effect of mismatch between the expected determiner and the perceived determiner represents the electrophysiological correlate of prediction.

In contrast, in the PMT, the feature pre-activation process triggered by the predictor (the picture) would be visual > lexical > sub-lexical[3]; in other words, full lexical access occurs earlier since it is not mediated by sub-lexical units. If the gender value is represented at the lexical level of processing it will trigger an earlier prediction update compared to a scenario in which gender is strictly sub-lexically encoded. Still, since gender can be reliably represented through both a lexical and a form-based route (Gollan & Frost, 2001), there would be no special need of activating sub-lexical representations in the PMT since gender values would be already lexically available. Crucially, our experiment did not only manipulate gender but also transparency. Let us consider the opaque words first. In order to predict the gender of opaque nouns participants need to rely on the lexical information, as there are no formal cues to give them clues about the gender. Therefore, in both groups of speakers we expect the unexpected determiner to elicit the same pre-activation timing as found in the previous experiments (the effect should be earlier in the PMT than in the WMT since the lexical information is earlier available in the former). The timing of opaque words represents our reference point for lexical prediction.

On the other hand, transparent words should display a different pattern in the two groups. In line with the previous experiment, we assume that Spanish (L1) speakers rely on lexical information for the pre-activation of gender, so we expect them to show a lexical effect (with similar timing to the opaque items) after they hear an unexpected article, both in the WMT and in the PMT.

In contrast, we predict Basque (L1) natives performing the WMT to present an earlier effect on the unexpected determiner (compared to Spanish natives), as they rely on sub-lexical information for the pre-activation of gender. When they see an image in the PMT, the lexical information will be accessed before the sub-lexical one. Thus, the gender would be pre-activated through the lexical route, but it would also be followed by an additional sub-lexical analysis. As a consequence, we expect Basque natives to show a delayed effect in comparison to Spanish natives.

---

3. *Depending on the theoretical model of reference, these representational dynamics can change a lot. For example, Roelofs (1992) suggests that after lexical-semantic access, an intermediate "lexeme" representation modulates sub-lexical phonological processing. In this framework, we mainly assume that (i) full lexical access effects should emerge earlier in the PMT compared to the WMT and that (ii) later effects imply lower level representations involving information about predictee inflection.*

If our expectations are correct, we will have further evidence that there are hierarchical levels of representation in language prediction; that prediction mechanisms can navigate this hierarchy quickly adjusting to the input; and that the language acquired first has an impact on language anticipation processes, even in early bilinguals with very high proficiency. This would add important information to language prediction research, and provide new insights into the mechanisms operating during monolingual and bilingual language processing.

3. Material and methods

3.1 Participants

Forty-two early bilinguals participated in the experiment. They were divided in two groups. The first group was formed by twenty-one native speakers of Spanish who were first exposed to Basque after the age of three (11 females; age range 19-29, mean: 23.38, SD: 3.24: Age of acquisition of Basque: 3.61 y.o., SD: 1.46).

Twenty-one native speakers of Basque (13 females; age range 18-33, mean: 25.66, SD: 5.45: Age of acquisition of Spanish: 4.23 y.o., SD: 1.33) who started to learn Spanish after the age of three formed the second group. All participants were right handed, their vision was normal or corrected to normal, and they had no history of neurological disorder. Before taking part in the experiment, participants signed an informed consent. They received a payment of 10 € per hour for their participation. The study was approved by the BCBL ethics committee.

In order to participate in the study, all the participants had to go through some language proficiency tests in both Spanish and Basque (results in Table 2). First, participants had to self-rate their language comprehension (on a scale from 1 to 10, where 10 was a native-like level; the result was averaged for speech comprehension, speech production, reading and writing). Basque speakers rated themselves very high in both Basque and Spanish, while Spanish speakers claimed they were better in Spanish than Basque.

Participants also went through a lexical decision task called LexTALE (Izura et al., 2014; Lemhofer & Broersma, 2012) to test their vocabulary knowledge. In Spanish, both groups showed the same very high score, but in Basque, Spanish speakers had lower scores than Basque natives.

Then participants had to name a sequence of pictures which increased in difficulty, in both languages. Again, both Spanish and Basque participants had the same native-range score in Spanish, but Basque speakers were better than Spanish speakers in Basque. Finally, all the participants were interviewed by balanced bilingual linguists who rated them on a scale from 0 to 5: in both languages no participants had a score below 4.

All the participants studied English at school, and claimed it was their third language. We applied the same measures to test subjects' English proficiency: there was no difference between groups, scores were good, but still much lower than Spanish and Basque (self-evaluation: 5.2, SD: 3.22; LexTALE, score 0−40: 22.43, SD: 5.15; picture naming: 44.07, SD: 7.22; interview: 3.88, SD: 1.44). Since participants were far more proficient in their two other languages than in English, we assumed that this third language could not influence the present design.

Table 2: General proficiency assessment of the participants in the two groups

Measure	**Spanish natives (N=21)**	**Basque natives (N=21)**
Self-evaluation (0-10)		
Spanish	9.37 (0.29)	9.39 (0.24)
Basque	8.02 (0.49)	9.04 (0.16)
LexTALE		
Spanish (0-60)	52.76 (5.21)	54.09 (4.13)
Basque (0-50)	34.38 (5.88)	46.04 (2.67)
Picture naming (0-65)		
Spanish	64.47 (0.92)	63.38 (1.62)
Basque	50.09 (9.63)	64.19 (1.47)
Interview (0-5)		
Spanish	5 (0)	5 (0)
Basque	4.14(0.47)	4.95(1.33)

3.2 Experimental design and materials

A list of 120 Spanish nouns was selected, where 60 nouns were transparent and 60 nouns were opaque (30 masculine and 30 feminine nouns per group). The transparent masculine nouns ended with "-*o*", which is the typical Spanish ending for masculine, (e.g. *cuchillo*, "knife [masc]"), while the feminine nouns had the feminine ending "-*a*" (e.g. *silla*, "chair [fem]"). Irregular nouns were excluded. Opaque nouns showed endings that were not informative of the grammatical gender (i.e., "-*e*", "-*n*", "-*l*", "-*s*", "-*j*", "-*r*", "-*d*", "-*z*").

The mean number of letters for transparent and opaque nouns was identical (mean: 5.76 letters, SD 1.51; range: 4–9 letters). In addition, transparent and opaque nouns did not differ for measures of concreteness (transparent nouns mean: 5.86, SD 0.58; opaque nouns mean: 5.67, SD 0.63); imageability (transparent nouns mean: 6.11, SD 0.73; opaque nouns mean: 6.09, SD 0.75) and familiarity (transparent nouns mean: 6.16, SD 0.42; opaque nouns mean: 6.12, SD 0.53) (EsPal, Duchon et al., 2013). Half of the nouns referred to artifacts, and half to natural objects. The words selected were also used to create the stimuli for the picture matching task. We found images that visually represented the nouns above. All the pictures were highly recognizable colour photographs (.png extension, white background, 2000x2000 pixels) obtained from online image collections. In order to be sure that a picture could only be related to one possible noun, we ran a naming test. Spanish-Basque bilinguals (N=20) who did not take part in the experiment saw 240 images, and named them with the first noun that came up to their minds. We only chose the images whose name was univocally expressed by all the 20 participants, and we came up with a final 120 pictures that could only represent our original list of nouns.

Both the words and the images were followed by an auditory noun phrase formed by a determiner followed by a noun. The determiner could be a definite article (*el* "the [masc]", *la* "the [fem]"), or a possessive adjective (*mi* "my", *su* "his, her") that was gender unmarked, hence neuter. The noun could either match or mismatch with the previously presented stimulus, but always matched with its own determiner.

The experiment had 5 conditions. There were 2 main conditions (that represented 53% of the trials). In the first one, both the determiner and the noun matched with the stimuli (written or visual); in the second, both the determiner and the noun mismatched with the stimuli. The gender value of the determined was balanced in the two conditions. In addition, we added a condition (26% of the trials) in which the determiner was neuter, and the noun matched with the stimuli. This last condition was introduced to reduce strategical prediction effects in our experimental design time-locked to the determiner. However, since the sound envelope of the neuter determiner was different compared to the experimental ones, we reasoned that the relative evoked response could not be directly compared with the experimental ones. Furthermore, there were 2 catch trials conditions (20% of the trials): in the first, a neuter determiner was followed by noun that mismatched with the previous stimulus; in the second, the determiner matched with the stimulus, but the noun did not.

For each word/image of the original list, 5 possible noun phrases were recorded, corresponding to each of the 5 conditions (e.g. *BOTELLA* “bottle [fem]”: *la botella* “the [fem] bottle”; *mi botella* “my bottle”; *el mando* “the [masc] remote control” ; *su mando* “his/her remote control”; *la corona* “the [fem] crown”).

Sound strings were recorded by a Spanish female speaker. Between the determiner and the noun there was a silence gap of about one second (range 1- 1.3s, the exact timing was measured for each item). All the items were checked for amplitude (recording, cuts, measures and standardization was done using Praat (Boersma & Weenink, 2007).

In order to assess how many milliseconds a listener needs to distinguish between the determiner *el* “the [masc]”, and the determiner *la* “the [fem]”, we ran a discrimination test with 10 participants who did not take part in the experiment. We took five NPs whose determiner was *el*, and five whose determiner was *la*. The audio files corresponding to the determiners were cut in order to create smaller time windows. Participants listened to audio fragments containing the first 60 ms; 70 ms; 80 ms or 90 ms of the determiner of each NP for a total of 40 trials, and they had to indicate (by spelling it aloud to an experimenter) whether they thought it was an *el* or a *la*. For the 60 ms bits the inaccuracy percentage was 54% for the determiner *el*, and 64% for the determiner *la*. Participants listening to audio files lasting 70 ms had an inaccuracy of 16% for *el*, and 28% for *la*. When audio fragments lasted 80 ms participants recognized all trials with the determiner *el*, but they got wrong 3% of the *la* determiners. There was 100% of accuracy when the audio files lasted 90 ms for both determiners. Therefore 90 ms is defined as the uniqueness point (Marslen-Wilson, 1987) of the determiner gender value, and will be later subtracted from the emerging ERP component timing.

For both the WMT, and the PMT, 3 lists were created. Each list had 240 visual stimuli, each of the 3 main conditions had 64 items, and the catch trial conditions had 24 items each. In each list, visual stimuli were repeated twice, but never in the same condition. Participants were never given the same list for both WMT and PMT, furthermore, the task sequence was alternated among participants, so that half of them first went through the WMT and then the PMT, and the other half did the opposite.

Within each list the words were balanced (all $p > 0.2$; based on EsPal, Duchon et al., 2013) for grammatical gender, word frequency (log-values: List1: 1.21, SD: 0.55; List 2: 1.26, SD: 0.52; List 3: 1.25, SD: 0.56), and number of letters (List1: 5.79, SD: 1.59; List 2: 5.79, SD: 1.46; List 3: 5.68, SD: 1.51). Finally, all the images were balanced among the lists, so that they were all equally distributed in all the conditions. No differences emerged between lists.

3.3 Procedure

The EEG experiment was run in a soundproof electrically shielded chamber with a dim light. Participants were seated in a chair, about sixty centimeters in front of a computer screen. Stimuli were delivered with the PsychoPysoftware (Peirce, 2007).

In the WMT, participants read words displayed in black letters on a white background. In the PMT, subjects saw images in the center of the screen. After a fixation cross (lasting 500 ms), the words or the images appeared on the screen for 350 ms. After the visual stimuli disappeared, they were immediately followed by auditory stimuli played by two speakers. The stimulus onset asynchrony (SOA) between visual and auditory stimulus was selected so that the onset of the auditory stimulus was time-locked to the on-going lexical processing of the visual word in the WMT and to sub-lexical processing in the PMT. In addition, studies employing cross-modal priming have shown that the effects tend to be more robust when the SOA is larger than 200 ms (Holcomb & Anderson, 1993). After participants heard the NP, they had to indicate whether the word they read or the image they saw matched with the noun they heard. The question appeared in the center of the screen as soon as the sound finished, and the subject could answer using the relative buttons on the keyboard: the response hand was counterbalanced across participants and list.

Numbers of correct responses were recorded, and RTs were calculated in milliseconds from the appearance of the question to the participant's key press. All trials were presented in a different random order for each participant.

WMT and PMT were presented in two parts of the same experimental session, with a small break between the two.

A brief practice session included five words in one session, and five images in the other, followed by the relative auditory stimuli and the yes-no questions. Participants were asked to stay still and to try to reduce blinking and eyes movement to minimum, especially during the auditory presentation. Overall, the experiment lasted one hour on average.

3.4 Electrophysiological recording and data analysis

EEG was recorded from 27 electrodes placed in an elastic cap (Easycap, www.easycap.de): Fp1, Fp2, F7, F8, F3, F4, FC5, FC6, FC1, FC2, T7, T8, C3, C4, CP5, CP6, CP1, CP2, P3, P4, P7, P8, O1, O2, Fz, Cz, Pz. All sites were online referenced to the left mastoid (A1). Additional external electrodes were placed on mastoids (A1, A2) and around the eyes (VEOL, VEOR, HEOL, HEOR) in order to detect blinks and eye movements. A forehead electrode served as the ground. Data were amplified (Brain Amp DC) with a bandwidth of 0.01−100 Hz, at a sampling rate of 250 Hz. The impedance of the scalp electrodes was kept below 5 kΩ, while the eye electrodes impedance was below 10 kΩ. Collected recordings were off-line re-referenced to the average activity of the two mastoids. Artifacts exceeding 100 μV in amplitude were rejected. Raw data were visually inspected and artifacts such as muscular activity and ocular artifacts were marked for subsequent rejection. On average, 7.3% of epochs were excluded as considered artifacts. There was no difference between conditions and groups in terms of artifact rejection.

For the analysis of the determiner, epochs of 1600 ms (from -600 ms to 1000 ms) were obtained, considering a -600 ms pre-stimulus baseline. For each condition, the average ERP waveforms were computed time-locked to the onset of the determiner. Epochs were averaged independently for each condition and subject.

In order to obtain a detailed exploration of the exact time course of the effects on the determiner, and define their evolution in time, we performed pairwise comparisons of the

ERP waveforms (match vs. mismatch) with point-by-point (one point every 4 ms) t-test for each electrode. We ran separate comparisons for each task, each gender type and each language group. To protect this analysis from false positives we employed the Guthrie and Buchwald (1991) correction that filters out effects which last less than 50 ms (12 consecutive time points) in less than three sensors. Importantly, this approach does not constrain the selection of the time interval of interest. However, we validated the relevant point-by-point effects that could reflect an interaction between the main factors of interest (match by transparency) within each Experiment (Word/Picture matching task) and within each group (Basque/Spanish natives) with further statistics. We selected the 100 ms-long time interval of interest and entered the average ERP activity – across the electrodes in which a significant effect emerged – in a three-way ANOVA with Match (two levels: match vs. mismatch), Transparency (two levels: transparent vs. opaque) and Electrode (variable number of levels depending on the electrodes of interest). P-values were Greenhouse-Geisser corrected.

For the analysis of the noun, we computed 3100 ms epochs (from -1600 ms to 1500 ms), applying the same pre-stimulus baseline used for the determiner, so that we could have all the NP electrophysiological time course, but the waveforms were time-locked at the exact onset of the noun for each condition. Here, the statistics time-locked to the noun focused on the time-interval classically related to lexical integration processing (300-700 ms, namely the N400).

4. Results

4.1 Behavioural

Inaccuracy (mean percentage of incorrect responses) and reaction times (RT) from accurate trials were analyzed.

In WMT, Spanish natives had a general average inaccuracy of 4.50 % (6.54% for the opaque words, and 2.46% for the transparent words). In Basque natives the average inaccuracy was 3.29 % (5.03% for opaque words, and 1.54% for the transparent words).

For the PMT, there was 3.11% of inaccuracy in Spanish natives (4.20% for the opaque words, and 2.02% for the transparent words). Basque natives gave 6.98% of inaccurate answers (8.01% for the opaque, and 5.95 for the transparent).

In the Spanish group, the mean reaction time for the WMT was 490 ms (SD: 0.37), while for the Basque groups it was 500 ms (SD: 0.30). In the PMT, Spanish natives has 490 ms (SD: 0.50) reaction time, and for the Basque natives it was 470 ms (SD: 0.24).

To analyze the reaction time values of the accurate trials for the WMT and the PMT we used two mixed-design ANOVAs, with Transparency and Condition as within factors, and with Group as between factor. No significant differences emerged from the analysis: there were neither main effects, nor interactions among variables (Table 2).

4.2 ERP data

4.2.1 Determiner

In order to better highlight the prediction effects time-locked to the determiner presentation, we report data in the time window until 1 sec. In the Figures below (1-4) we report the ERPs for the electrodes in the left hemisphere scalp region in which significant effects emerged; the point-by-point analyses showing activation timing across all the electrodes; and the scalp topography relative to the 250-350 ms and 400-500 ms time windows.

In the WMT, for the Spanish group, determiners preceding the opaque nouns display a more negative effect for Mismatch items, likely reflecting a N400 modulation: this long-lasting effect starts at about 420 ms (330 ms post-uniqueness point), with a widespread distribution in all the electrodes. Basque natives showed a similar short-living negative effect starting 430 ms (340 ms post-uniqueness point).

The WMT relative to the determiners preceding transparent nouns showed different results. Spanish natives display a strong effect of Prediction in the time interval starting around 380 ms (290 ms post-uniqueness point): the effect is more negative for Mismatch compared to Match condition and it could be interpreted as an N400 evident in frontal and parietal electrodes.

Crucially, in the same condition Basque natives have an earlier effect in the time window from 240 ms to 350 ms (150 ms post-uniqueness point). The effect is most positive for the Mismatch condition and particularly evident on the left parietal electrodes, and it can be interpreted as a P200 effect.

The early effect is followed by an N400 effect starting at about 430 ms (340 ms post-uniqueness point) with a wider scalp distribution.

The three-way ANOVA run to check for robustness of Prediction effect in the 250-350 ms time window (for the electrodes P7, P3, and CP5) in both groups showed an interaction between Match (congruency between stimulus and determiner) and Transparency only for Basque natives [Spanish natives: $F(1,20)=0.63$, $p>0.4$, ges=0.002 ; Basque natives: $F(1,20)=5.01$, $p<0.05$, ges=0.006]; importantly, it also showed an interaction between Match and Group [$F(1,40)=3.89$ $p<0.05$, ges=0.012].

The results of the PMT display a different pattern. For determiners preceding opaque nouns, in both Spanish and Basque natives there is a robust effect in an early time window, from 250 (160 ms post-uniqueness point) to 350 ms. The effect is more positive for mismatching items compared to matching items. We assume this to be an early lexical effect.

The effect on transparent nouns was different. In Spanish natives, determiners preceding transparent nouns elicit an increased negative effect for mismatching items at about 420 ms (330 ms post-uniqueness point). The effect is frontally distributed. A more left-posteriorly distributed negative effect for mismatching items is also present in Basque natives: it starts later, at 550 ms (460 ms post-uniqueness point).

For the 250-350 ms time window, in the Spanish group, a three-way ANOVA showed interaction between Match and Transparency in the electrodes of interest [$F (1,20) = 10.51$, $p < 0.01$, ges = 0.011]. The Basque natives displayed the same interaction [$F (1,20) = 4.65$, $p < 0.05$, ges = 0.006]. In the later time intervals, we ran two separate ANOVAs considering Group and Match as separate factors. In the 550-650 ms time interval we observed a significant interaction between the two factors when considering the electrodes showing the significant difference for transparent items in the Basque group [$F (1,40) = 3.60$, $p < 0.05$, ges = 0.005]. However, no interaction was observed in the 400-500 ms interval (in contrast to the electrodes showing the significant difference for transparent items in the Spanish group) [$F (1,40) = 1.66$, $p > 0.1$, ges = 0.001]. We can thus conclude that Spanish natives showed a statistically reliable prediction effect in the 400-500 ms that was not strong enough to be reliable for Basques. In the later 550-650 ms time interval, a reliable prediction effect was observed in the Basque native group but not in the Spanish native group.

4.2.2 Noun

ERP results on the onset of the noun are clear-cut and straightforward. Given the long duration of the effect we found, the time window taken into consideration for the analysis is longer ending at 1.5 sec.

In both groups, both WMT and PMT produced a strong mismatch effect starting at about 300 ms and ending at about 800 ms. No significant differences were found between transparent nouns and opaque nouns. We assume this result to reflect a semantic integration effect.

We ran a three-way ANOVA with Match, Hemisphere and Longitude as factors, for the 300-700 ms time window and for all the electrodes. The analysis confirmed the results: the main effect of Match and the interactions between Match and Hemisphere and Match and Longitude are all significant (Table 3).

Measure	WMT	PMT
Condition	F (4, 52) = 1.36, p = .251, η_p^2 = .43	F (4, 52) = .62, p = .649, η_p^2 = .16
Transparency	F (1, 38) = .76, p = .389, η_p^2 =.20	F (1, 38) = .29, p = .596, η_p^2<.01
Group	F (1, 38) = 1.36, p = .758, η_p^2 = .03	F (1, 38) = .33, p = .567, η_p^2<.01
Condition*Group	F (4, 52) = 1.32, p = .265, η_p^2 = .37	F (4, 52) = .79, p = .531, η_p^2 = .20
Transparency*Group	F (1, 38) = 2.16, p = .150, η_p^2 = .54	F (1, 38) = 89, p = .353, η_p^2 = .23
Condition*Transparency	F (4, 52) = 1.34, p = .260, η_p^2 = .24	F (4, 52) = .54, p = .706, η_p^2 = .14
Condition*Transparency*Group	F (4, 52) = .31, p = .869, η_p^2<.01	F (4, 52) = .60, p = .662, η_p^2 = .01

Table 3: ANOVA results of WMT and PMT reaction time values

	WMT		**PMT**	
	Spanish	**Basque**	**Spanish**	**Basque**
Transparent	F(1,20)= 32,00, p< 0,05, ges=0,25	F(1,20)= 34,14, p< 0,05, ges= 0,16	F(1,20)= 30,25, p< 0,05, ges= 0,21	F(1,20)= 43,92, p< 0,05, ges= 0,21
Opaque	F(1,20)= 40,70, p< 0,05, ges= 0,28	F(1,20)= 27,12, p< 0,05, ges= 0,18	F(1,20)= 71,24, p< 0,05, ges= 0,31	F(1,20)= 26,54, p< 0,05, ges= 0,18

Table 4: ANOVA main effects of Match on the noun (300-700 ms)

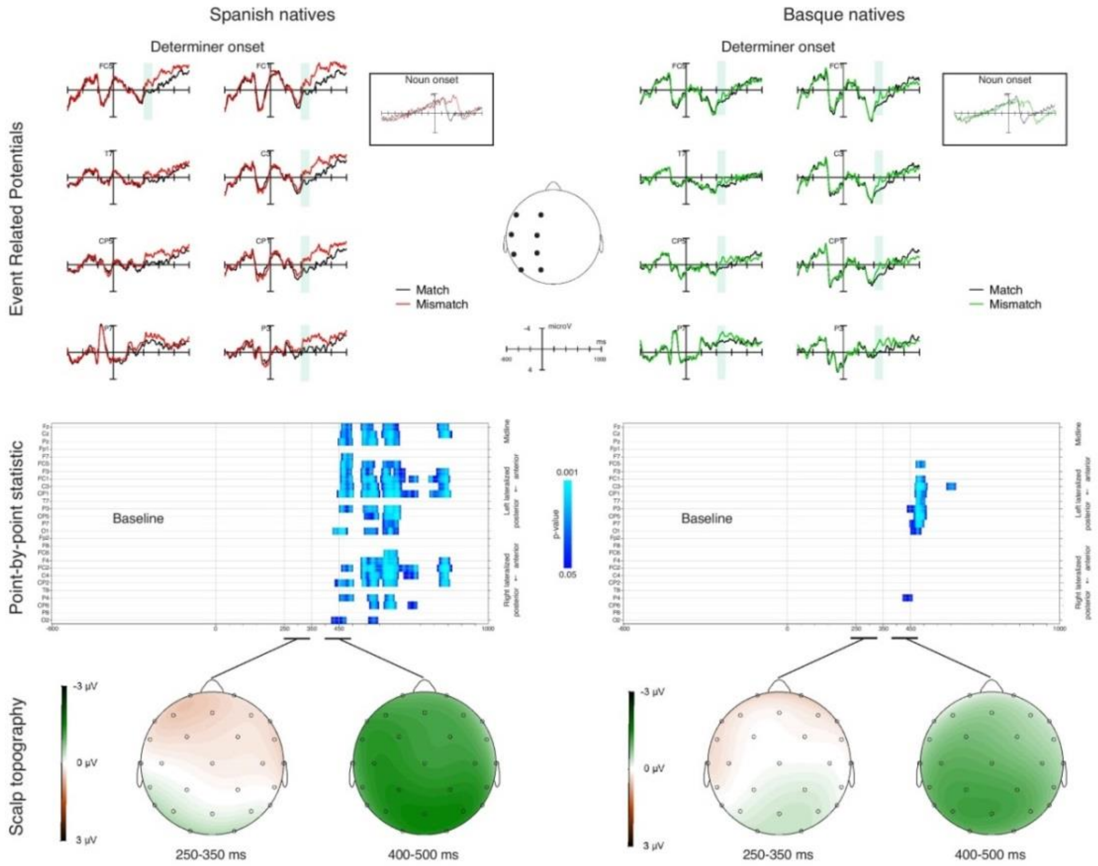


*Figure 8: ERPs for the **WMT** relative to the match and mismatch condition (Prediction effect) time-locked to the presentation of the determiner preceding the predicted **opaque** noun (plotted in -600 to 1000 for better display of the ERP effects). We separated the two groups of participants and plotted all the electrodes. Shadowed differences are the statistically significant ones relative to the 250-350 ms time window. The rectangles on the right side of the graphs present the grand average ERPs relative to the match and mismatch condition (Integration effect) time-locked to the presentation of the **opaque** noun (plotted in -1600 ms to 1500 ms), for all the electrodes. The point-by-point plot analysis of variance for each electrode (Guthrie and Buchwald, 1991, corrected) is presented below. We report the main effect of the Prediction factor: vertical blue lines indicate the interval in which the effect emerged. In the lower panels we present the topographical distribution of the difference effect (mismatch minus match) in the two time intervals of interest (250-350 ms; 400-500 ms).*

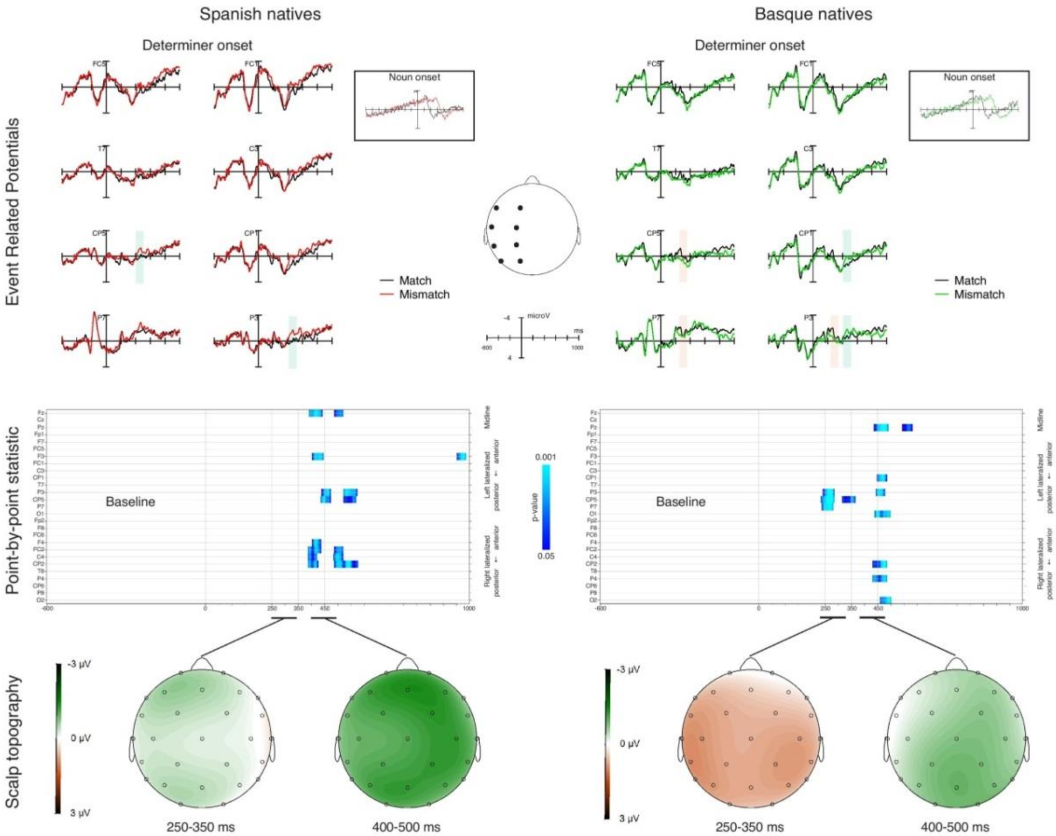


*Figure 9: ERPs for the **WMT** relative to the match and mismatch condition (Prediction effect) time-locked to the presentation of the determiner preceding the predicted **transparent** noun (plotted in -600 to 1000 for better display of the ERP effects). We separated the two groups of participants and plotted all the electrodes. Shadowed differences are the statistically significant ones relative to the 250-350 ms time window. The rectangles on the right side of the graphs present the grand average ERPs relative to the match and mismatch condition (Integration effect) time-locked to the presentation of the **transparent** noun (plotted in -1600 ms to 1500 ms), for all the electrodes.The point-by-point plot analysis of variance for each electrode (Guthrie and Buchwald, 1991, corrected) is presented below. We report the main effect of the Prediction factor: vertical blue lines indicate the interval in which the effect emerged. In the lower panels we present the topographical distribution of the difference effect (mismatch minus match) in the two time intervals of interest (250-350 ms; 400-500 ms).*

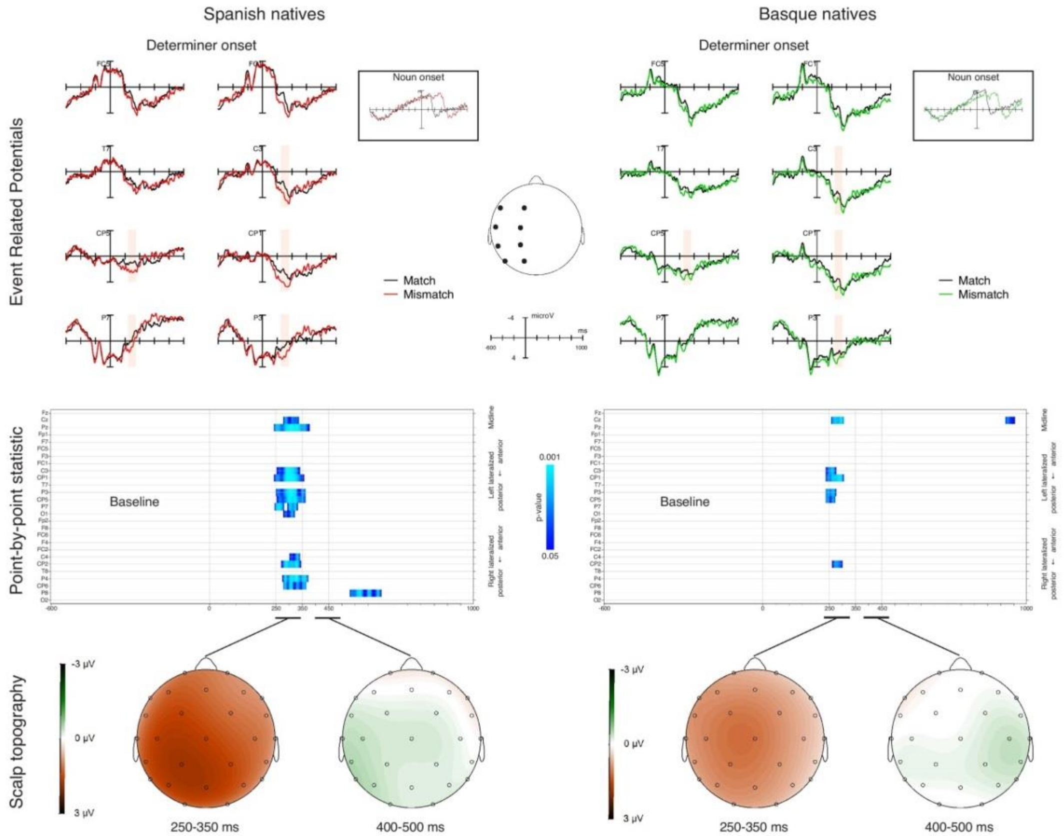


*Figure 10: ERPs for the **PMT** relative to the match and mismatch condition (Prediction effect) time-locked to the presentation of the determiner preceding the predicted **opaque** noun (plotted in -600 to 1000 for better display of the ERP effects). We separated the two groups of participants and plotted all the electrodes. Shadowed differences are the statistically significant ones relative to the 250-350 ms time window. The rectangles on the right side of the graphs present the grand average ERPs relative to the match and mismatch condition (Integration effect) time-locked to the presentation of the **opaque** noun (plotted in -1600 ms to 1500 ms), for all the electrodes. The point-by-point plot analysis of variance for each electrode (Guthrie and Buchwald, 1991, corrected) is presented below. We report the main effect of the Prediction factor: vertical blue lines indicate the interval in which the effect emerged. In the lower panels we present the topographical distribution of the difference effect (mismatch minus match) in the two time intervals of interest (250-350 ms; 400-500 ms).*

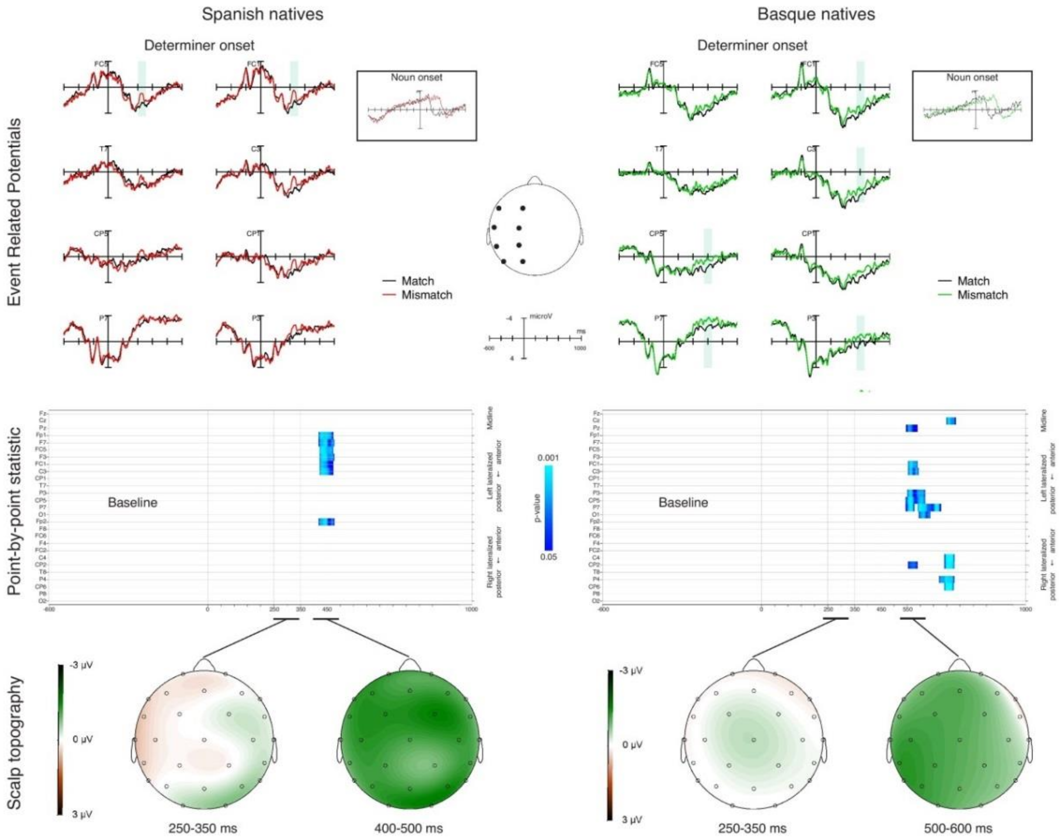


*Figure 11: ERPs for the **PMT** relative to the match and mismatch condition (Prediction effect) time-locked to the presentation of the determiner preceding the predicted **transparent** noun (plotted in -600 to 1000 for better display of the ERP effects). We separated the two groups of participants and plotted all the electrodes. Shadowed differences are the statistically significant ones relative to the 250-350 ms time window. The rectangles on the right side of the graphs present the grand average ERPs relative to the match and mismatch condition (Integration effect) time-locked to the presentation of the **transparent** noun (plotted in -1600 ms to 1500 ms), for all the electrodes. The point-by-point plot analysis of variance for each electrode (Guthrie and Buchwald, 1991, corrected) is presented below. We report the main effect of the Prediction factor: vertical blue lines indicate the interval in which the effect emerged. In the lower panels we present the topographical distribution of the difference effect (mismatch minus match) in the two time intervals of interest (250-350 ms; 400-500 ms for the Spanish natives and 500-600 ms for Basque natives).*

5. Discussion

The goal of the present study was to provide evidence for the existence of hierarchical levels of representation in language prediction.

For this purpose, we used two groups of highly proficient bilinguals going through a WMT and a PMT. Subjects read a word, or saw a picture on a screen, followed by an auditory NP where the noun could be congruent or not with the preceding visual stimuli, and the determiner agreed with the noun. Also, the noun could be transparent or opaque. The ERP effects recorded on the determiner preceding the noun (expected vs. unexpected determiner) revealed the time course of the processes operating inside linguistic prediction. In the WMT, for the unexpected article preceding opaque words we found the same pre-activation timing for both Spanish and Basque natives. The effect started at ~430 ms (340 ms post-uniqueness point) in both groups, and it lasted longer in the Spanish group. This effect can be interpreted as an N400. The result is in line with our expectations: participants have to rely on their lexical knowledge (as reflected by the ERP lexical effect) for the gender prediction of opaque words, which do not provide formal cues about their gender. The effects on determiners preceding unexpected transparent words showed different results. Spanish native ERPs revealed a negativity at ~390 ms (300 ms post-uniqueness point) roughly similar to the opaque words. This lexical effect was observed in Basque natives too, but crucially, they also displayed an early prediction effect at ~240 ms (150 ms post-uniqueness point). This confirmed our expectations: we had assumed that Basque natives rely more on sub-lexical, word-form related information for gender prediction of transparent words and we were expecting this prediction process to be reflected by the modulation of an early ERP component related to attentional perception, like the ~240 ms positivity found here. On the other hand, in Spanish natives predicting transparent words, we were expecting to see a lexical ERP effect (similar to the opaque words) as they would rely more on the lexical information for gender prediction, irrespective of the transparency of the target noun.

In the PMT, we assumed the lexical information to be available earlier compared to the WMT, as lexical access triggered by the picture would not be mediated by the sub-lexical units of the word, and would be direct. For the unexpected determiners preceding opaque words we found a similar effect at ~250 ms (160 ms post-uniqueness point) for both

Spanish natives and Basque natives. We assume that this effect reflects a lexical/semantic process (if we assume the prediction/production parallelism, this “lexical timing” is typically reported in picture naming studies: Indefrey & Levelt, 2004; Strijkers & Costa, 2016). Here our expectations were confirmed, as the gender of opaque words can only be accessed and predicted through lexical information.

As predicted, the ERPs recorded on determiners preceding transparent nouns displayed an effect at ~430 ms (340 ms post-uniqueness point) in the Spanish group, and later effect (~550 ms, 460 ms post-uniqueness point) in the Basque group. Basque natives presumably performed additional sub-lexical analysis after the lexical one, resulting in later timing. Besides, in comparison to the determiners preceding opaque words, the effect happens later in time in both groups: a possible explanation for this dissimilarity will be given later in the discussion.

5.1 The role of native experience

This experiment provides further evidence that lexical prediction does not involve predicting only the semantic traits of a word (Federmeier, 2007, Federmeier et al., 2002), but also its grammatical features. These features can be predicted even when they are not present in the bilinguals’ L1, in support of the idea that there are no separate domains of prediction, but that prediction is a unified mechanism looking for informative cues independently of the language processed.

Nonetheless, a central point emerging from the study is that language prediction is tuned to the characteristics of the L1 and consequently is extremely sensitive to the native experience.

As the only difference between groups is the language learned first, we can assess that the prediction differences found between Basque and Spanish natives are due to the early L1exposure. During the first three years of their lives (in which they were only exposed to the L1) Basque speakers developed their prediction mechanisms following the regularities of Basque language. These mechanisms still influence Basque natives' prediction processes even if they have been daily exposed to the statistical regularities of Spanish all the rest of their lives. On the other hand, speakers originally tuned to Spanish process transparent and opaque items equally during prediction, as a result of the more robust mapping of grammatical gender from lexical information in this language. This had already been found to be true not only in experiments with words in isolation (Caffarra et al., 2014), but especially in experiments investigating the effect of transparency during determiner-noun agreement (Caffarra & Barber, 2015).

It is worthy of notice that the difference in prediction mechanisms between the two groups cannot be due to a reduced proficiency for Basque natives in Spanish, since in the general proficiency assessment the two groups were perfectly balanced in this language, and also the behavioral results in both tasks did not reveal any significant differences.

5.2 Hierarchical levels of prediction

More importantly, this study provides stronger evidence for the existence of cognitive hierarchies in language prediction. Basque natives' neurophysiological responses display two levels of language prediction in the WMT. At one level they rely on the sub-lexical information of the word that is going to come next, in order to predict its gender. At another stage, gender can be predicted according to the lexical information they have about that word (as stored in the mental lexicon). The two levels are differently represented in terms of neural timing, with the first operating earlier, at about 150 ms after gender disambiguation of the determiner, and the second operating later, in the N400 time window, as clearly emerged from the WMT. This result supports what we recently observed during sentence reading (Molinaro et al., 2016; under review), where both groups showed a lexical N400-like effect for lexical prediction during sentence reading, while the Basque natives showed an earlier N200-like effect only for the pre-activation of transparent words. Also, we have replicated this outcome using a completely different task, as we did not use highly constraining contexts in order to elicit prediction, as in the previous experiment, but we used a WMT where prediction was evoked by a single word on a screen. This is evidence that the kind of representation pre-activated in both Molinaro et al. (2016; under review) and in the WMT is strictly word-related. In addition, the prediction effects that we found are observable independently of the kind of experimental setting, supporting the idea that prediction-mechanisms are able to quickly and flexibly adjust to the different tasks a language user has to go through. Furthermore, in the present study, the predictees were not visual but were audio, providing critical evidence that prediction mechanisms are modality independent. The existence of hierarchies clearly emerges from the WMT, but what can the PMT tell us about them?

5.3 An interface where hierarchical levels communicate

Given the characteristics of the PMT, in our opinion, a good way to approach and discuss the PMT results would be through psycholinguistic production models. According to these, the mental lexicon is conceived as a network in which concept, lemma (word stems) and lexeme (morphological units) are represented as spreading-activated, sequential, independent nodes (Roelofs, 1992; Levelt et al., 1998). It is feasible to assume that, in order to perform the PMT (and ultimately predict the word that is going to come with its relative gender), participants exposed to a picture activate conceptual/semantic information, which is directly followed by the activation of the lemma level representation; lastly they activate the information at the lexeme level.

In the PMT, unexpected determiners preceding opaque words elicited a lexical ERP effect at ~160 ms in both Spanish and Basque natives. On the other hand, unexpected determiners preceding transparent words produced a much later ERP effect, at about 330 ms for Spanish natives, and ~460 for Basque natives.

The hierarchical spreading activation dynamics would be the same for both transparent and opaque words. Nonetheless, the two groups respond differently in the two cases, with the opaque words showing a much earlier activation compared to the transparent words. Why do the two groups display a difference in the prediction time course of transparent and opaque words? Also, why is there a difference in timing between groups only on determiners preceding transparent words?

One way to explain the different timing would be to suppose that the prediction of opaque words is based on the lemma representation participants have about those words, while for the prediction of transparent words participants rely more on the lexeme representation, so they predict the former earlier than the latter, as the former are activated before, in a cascade process.

A similar way to account for the difference found in the PMT between the prediction of transparent and opaque words would be to assume the existence of an interface between the lemma level and the lexeme level. The “160 ms effect” found on the determiner preceding the opaque words is a fully lexical effect. It could not be anything else because the gender of opaque words can only be extracted and predicted through the lexical route. In contrast, the later effect emerging at 340 ms in Spanish natives, and at 460 ms in Basque natives for articles preceding transparent words would be due to additional processing resources at a lower hierarchical level. We think that the gender of the transparent words cannot be predicted only with the lexical information, but needs (or simply is given) more support from the sub-lexical information level. This information is not directly available because of the nature of the PMT, and the resulting effect reflects an operation going on at the interface between lemma-related and lexeme, inflection-related information (visual > lexical > sub-lexical). This hypothesis is supported by the timing of the effect that, based on production models (Indefrey & Levelt, 2004; Strijkers & Costa, 2016), would reflect activation of sub-lexical phonological information.

Further support for the lemma-lexeme interface is given by the difference in timing that emerged on determiners preceding the transparent words, between Spanish and Basque natives. In the conflict at the interface level between the lemma features activated by the picture and the lexeme features accessed after it, Basque natives would rely more on sub-lexical analysis for the extraction and prediction of gender, while Spanish natives would depend on the support of the lexicon to perform the prediction. This process is reflected by an effect happening ~130 ms later than that of the Spanish natives.

The existence of an interface between lexical and sub-lexical knowledge that permits interaction between the two kinds of process might be partially visible also in the WMT. In this task we assumed the processing sequence to be visual > sub-lexical > lexical. As indicated above, the Spanish natives display a lexical effect at 290 ms on determiners preceding transparent words; this effect is also present for the determiner preceding opaque words in both groups, but it happens later, at 340 ms.

It is possible that the earlier lexical effect found in Spanish natives predicting transparent words in the WMT is due to the interface interaction between lexeme and lemma analysis. This interaction does not take place for the prediction of opaque words because their gender cannot be extracted through the sub-lexical information, and it does not happen in Basque natives predicting transparent words, as the first analysis they run is already fully sub-lexical. Unfortunately, the ANOVA looking for interactions between transparency, match and group did not show any significant value in this time window (290-340 ms), therefore there is no strong evidence supporting this idea.

Future studies can possibly shed light on the existence of the interface between lemma-related and lexeme-related information in language prediction.

5.4 Results on the Noun

Finally, the ERPs recorded on the incongruent noun showed a strong late N400 effect. The effect was similar in both groups, in both tasks, in all the conditions. There was no significant difference and no interactions among the variables. This lexical effect is coherent with several previous studies on bilingual processes (Hahne, 2001; Hahne & Friederici 2001; Weber-Fox & Neville, 1996), and confirms that integration semantic processes are not influenced by the native language experience, especially when proficiency is very high. Also, this outcome is in line with the behavioural results, where no difference between groups was found.

Given the speakers' capacity to mediate between native language experience, new language environment cues and task requirements, we assume prediction mechanisms to be extremely plastic and flexible.

Conclusion

With the present set of data, we would like to emphasise the plasticity of language prediction mechanisms. This study not only shows that predictions can rapidly adapt to new contexts, selecting all the available cues in order to construct an internal representation of the new environment, but also that early experience models the way in which different cues are weighted to obtain optimal predictions. The whole process is extremely flexible, with neural timing constantly ready to adapt to the environmental demands and to mediate between native experience settings and new language feature requirements.

Chapter 4

The influence of Language Prediction on Perception

1. Testing language interference on prediction

In the previous experiments we provided evidence that language works in a predictive fashion: users are able to anticipate not only the item that is coming next, but also its features. As emerged from our ERP studies, prediction mechanisms are detectable extremely early in the time course of language comprehension. We highlighted the role of context and world knowledge on language processing, assuming that prediction is a top-down mechanism. Also, in the previous studies we progressively reduced the amount of context participants could base their predictions on (from constraining sentences to single words) and still found substantial top-down anticipative effects.

Both the experiments provided evidence for the existence of hierarchical levels of language prediction. Prior expectations about the world facilitate language processing: every level of the hierarchically organized system that constitutes the brain works to predict the activity in the level below.

Nonetheless, another clear pattern emerged from the previous studies. In early bilinguals, language prediction appears to be tuned to the regularities of the L1, and native exposure seems to have a fundamental role in shaping prediction mechanisms in participants. Crucially, in both experiments, Basque natives displayed a processing advantage in the anticipation of Spanish nouns, compared to Spanish natives. The advantage was visible in the activation hierarchy, and in particular in the capacity to rely on both the sub-lexical information and on the lexical information (in case the first did not succeed) for the prediction of the gender of the nouns. Spanish natives instead could only count on the lexical information, given their robust mapping of grammatical gender at a lexical level. We supposed that the Basque natives' advantage is due to the characteristics of the Basque language, and to interaction of their L1 with the L2. We also found some evidence that the two language systems can communicate at an interface level, and plastically influence each other.

Another possible explanation for the Basque natives' advantage in lexical prediction could be that they were better predictors simply because they were more balanced between languages compared to the Spanish natives. In fact, Basque natives' proficiency in Spanish was higher than Spanish native's proficiency in Basque.

The present experiment is a control study aiming at testing the possibility that the presence of the L1 can actually have a negative interference on tasks performed in the L2. In order to do this, we had to eliminate the factor we think caused the L2 advantage in the previous studies, namely the morphology itself. We bypassed morphology by exploring a different, faster path: visual perception.

2. Can language influence perception?

How far does language prediction go? Can language prediction be so fast to affect and change perception itself? Ultimately, is it possible for language to alter and shape cognition?

In this chapter we will explore the possibility that language not only influences reasoning, memory, and "high level" cognitive functions, but has effects on "lower-level" processes.

In this paragraph the debate on the relationship between language, perception and cognition will be briefly introduced.

The knowledge of what something looks like can be activated in many ways. For instance, after learning that cats meow, hearing a meow can activate the corresponding visual knowledge, as showed by facilitated visual recognition and category discrimination tasks (Lupyan & Thompson-Schill, 2012; Edmiston & Lupyan, 2013). Another way in which visual knowledge can be activated is through words. However, unlike other perceptual cues, words are categorical, meaning that the word "cat" can be used to refer to any cat. Words activate mental states corresponding to general categories.

Many researchers believe that such verbal activation of knowledge leaves perceptual processing unchanged (Gleitman & Papafragou, 2005; Klemfuss et al., 2012; Firestone & Scholl, 2014), and consider words as "pointers" to high-level conceptual representations (Bloom, 2000; Jackendoff, 2002; Li et al., 2009). In opposition to this view, other researchers think that words can affect visual processing with the effect of altering how incoming information is processed from the very start (Thierry et al., 2009; Lupyan, 2012a; Boutonnet et al., 2013; Lupyan & Ward, 2013; Francken et al., 2015; Kok & de Lange, 2014; Kok et al., 2014).

Behavioral experiments measured the effects of verbal labels on the activation of conceptual information in opposition to nonverbal sensory cues (Lupyan & Thompson-Schill, 2012; Edmiston & Lupyan, 2013). Participants completed a category verification task in which an auditory cue, which could be either a spoken category label (e.g., "cat") or a natural sound (e.g., "meow") preceded a visual image. Participants had to indicate if each cue-image pair matched by pressing the *yes* or *no* keys. The logic beneath the experiments was that a response to a visual stimulus can be altered by a cue presented prior to the target stimulus. From the results it emerged that words were more effective at activating the concept, as evidenced by significantly shorter reaction times, compared to sound trials. In other words, people were faster in recognizing a picture of a cat after hearing "cat" than after hearing a meow.

These findings suggested that a concept activated via different means, for example, via an auditory verbal label or via nonverbal auditory information, is detectably different. Specifically, the visual aspects of a category (e.g., the shape of a cat) are more effectively activated after hearing a word rather than a nonverbal sound. The authors speculated that labels modulate ongoing perceptual processing, therefore the cat that one sees after hearing the word "cat" is, somehow, not the same cat as the cat one sees after hearing a meowing sound or just thinking about cats. This is consequence of the fact that visual representations are subject to robust top-down modulations and language is a form of such top-down modulation.

These studies suggest that words are not simply a "pointer" or a means to access a nonverbal concept. Rather, they provide a special way to activate the representation that constitutes the concept.

In order to further test the hypothesis that words are able to provide top-down guidance at the very earliest stage of visual processing, and measure the timing of this effect, Boutonnet & Lupyan (2015) created an elegant ERP experiment. Following the previous studies, they recorded brain EEGs while participants indicated whether a picture matched a previously presented word (e.g., "cat") or an equally informative but nonverbal cue (e.g., cat meow). The idea was that if labels are especially effective at triggering visual representations and give to the visual system a set of priors, processing an image in light of these priors should help to accept congruent images and reject incongruent ones (Kok et al., 2014). Therefore, the difference in recognizing an image when its category is cued verbally versus nonverbally should be reflected in changes to very early ERP signals such as the P100, a component known to reflect processing of low-level visual features and classically associated with bottom-up processes (Mangun & Hillyard, 1991; Allison et al., 1999). If instead the advantage has an exclusively a semantic locus, then differences in brain activity elicited by label-cued versus sound-cued pictures would occur late in time and be only reflected by ERP components commonly associated with semantic integration, namely the N400 (Kutas & Federmeier, 2011).

Their results unequivocally supported the hypothesis that the word advantage has a perceptual locus. Verbal labels elicited more positive and earlier P100s than nonverbal stimuli (in the posterior electrodes). This suggests that after hearing a word, the brain processes responsible for generating the P100 were already sensitive to the object's category. Importantly, the P100 was sensitive to the congruence between the cue and the target, but only when participants were cued by a word, meaning that the brain was already sensitive to the object's category before perceiving the word. In addition, incongruent cues elicited greater N400s compared with congruent ones (a classic N400 semantic incongruity congruity effect; Kutas & Hillyard, 1984; Kutas & Federmeier, 2011), these effects were equivalent for verbal and nonverbal cues, supporting the fact that the two stimulus types were equivalently unambiguous at a semantic level.

These findings are explicable in the light of predictive processing frameworks. The stimulus that participants hear before the picture activates visual predictions. For instance, the label "cat" or the meowing sound activates visual representations corresponding to a cat shape. The picture (target) that successively appears is processed in light of these predictions. If the predictions are accurate, they will help to recognize an image from the cued category or reject an image from a non-matching category. What distinguishes the two cue types is that words denote categories, while sounds do not. The word "cat" denotes all cats, abstracting the peculiarities of particular exemplars, in contrast, nonverbal cues, are necessarily linked to a particular cat, making them less categorical. This outcome has been found to be correct for native speakers, but what about our bilingual speakers? Would bilinguals' visual system be advantaged as well by verbal labels in the L1? Or would their L1 interfere with the L2 prediction?

2. The present study

We conceived this experiment as a control study. It aims at understanding whether highly proficient early bilinguals are able to activate visual predictions like native speakers do. Also, the experiment has the goal to further analyze the interaction and/or interference between languages (if any) at the early visual prediction processing stage.

In order to compare this experiment to the previous ones we analyzed two similar groups of bilinguals: participants were Spanish (L1) - Basque (L2) speakers, and Basque (L1)-Spanish (L2) speakers. Like in the previous two studies, the only difference between the two groups was that the second language had been learnt at the age of three years old in both groups.

In this experiment, participants heard a verbal cue in Spanish (e.g., *perro,* "dog") or an equally informative non-verbal cue (e.g., dog bark). After the auditory stimuli, a picture appeared on a screen. The picture could match or mismatch with the auditory cue. Participants simply had to indicate whether the picture matched or not with what they heard by using two keys on the keyboard. We measured ERPs time-locked to the onset of the image.

According to the experiments presented above, native speakers are faster and better at recognizing a picture of a dog after hearing "dog" than after hearing a bark. In this study, the label advantage should provide the visual system with a set of priors that bias the processing of incoming stimuli. Therefore, in Spanish native speakers performing the task in Spanish, we expected differences in recognizing an image when its category is cued verbally versus non-verbally to be reflected in changes to early electrophysiological signals like the P100. This outcome would replicate Boutonnet & Lupyan (2015).

Importantly, in our study we will have the possibility to analyze a group of non-native high proficient speakers. As mentioned before, from the two experiments presented in the previous chapters emerged a processing advantage in Basque (L1) -Spanish (L2) speakers in comparison with the Spanish (L1) - Basque (L2) speakers. The advantage was reflected by an early ERP activation when participants had to predict the gender of the Spanish nouns. We assumed this advantage to be related to the characteristics of the Basque language, to which they rely in order to predict upcoming Spanish words. Importantly, in this last study, morphology is not taken into consideration, as we are evaluating language labels (L2 words in isolation) as opposed to non-verbal sounds.

If the advantage emerged before is related to the participants L1 typological characteristics interacting with the L2 prediction processes, here we should not find any advantage in Basque speakers. The “morphological” advantage would not be available. As a consequence, Basque natives should not be faster nor better than Spanish natives in recognizing a picture after a word.

Once eliminated this factor, we can safely analyze “pure” visual prediction mechanisms in bilingual speakers. The results relative to the Basque natives may display different outcomes. In the first case, Basque natives might not show any difference in respect to the Spanish natives in recognizing pictures, as a consequence of the fact that they are highly proficient in the L2. A similar P100 effect to the one elicited in Spanish natives should emerge on the image. This would mean that they are perfectly balanced speakers, and there is no difference for them between predicting in L1 or in L2, even at the fastest stages of anticipation like visual prediction.

A second option would be that Basque natives are not able to generate visual predictions like Spanish natives do. This would result in an absence of P100 effect in Basque natives when presented with the image after hearing the verbal label. Words would still be presented in their L2, and this could be reflected in a slow-down in processing, compromising very fast mechanisms like visual recognition. In addition, the very fact that Basque speakers are very balanced bilinguals could cause interference from L1 onto the L2, and compromise L2 visual prediction processes.

Finally, in line with the previous studies, from the analysis on non-visual semantic representations (to which the images are matched during recognition), we expect both groups to show the same effect. A strong N400 effect should be elicited by both verbal and non-verbal cues and found on the images that do not match with the cues in opposition to those who match.

3. Material and methods

3.1 Participants

We tested forty-four early bilingual speakers. They were divided in two groups. The first group was composed of twenty-two native speakers of Spanish who were first exposed to Basque after the age of three (12 females; age range 19-29, mean: 23.38, SD: 3.24: Age of acquisition of Basque: 3.61 y.o., SD: 1.46).

The second group was formed by twenty-two native speakers of Basque (13 females; age range 18-33, mean: 25.66, SD: 5.45: Age of acquisition of Spanish: 4.23 y.o., SD: 1.33) who started the acquisition of Spanish after the age of three. All participants were right handed, with no history of neurological disorder. Their vision was normal or corrected to normal. Also, they received a payment of 10 € per hour for their participation. Before taking part in the experiment, participants signed an informed consent. The study was approved by the BCBL ethics committee.

Before taking part to the study, all the participants had to go through a few language proficiency tests in both Spanish and Basque (statistical results in Table 5). Participants had to self-rate their language comprehension (on a scale from 1 to 10, where 10 was a native-like level; the result was averaged for comprehension, production, reading and writing). Basque speakers rated themselves very high in both Basque and Spanish, while Spanish speakers said they were better in Spanish than Basque. Participants also performed the "LexTALE", a lexical decision task (Izura et al., 2014; Lemhofer & Broersma, 2012) that tested their vocabulary knowledge. In Spanish, both groups displayed the same very high score, but in Basque, Spanish speakers had lower scores than Basque natives. In addition, participants had to name in both languages a series of pictures which increased in difficulty. Here as well, both Spanish and Basque participants had the same native-range score in Spanish, but Basque speakers were better than Spanish speakers in Basque. Finally, all the participants were interviewed by balanced bilingual linguists who rated them on a scale from 0 to 5: in both languages no participants had a score below 4.

Measure	Spanish natives (N=22)	Basque natives (N=22)
Self-evaluation (0-10)		
Spanish	9.37 (0.29)	9.39 (0.24)
Basque	8.02 (0.49)	9.04 (0.16)
LexTALE		
Spanish (0-60)	52.76 (5.21)	54.09 (4.13)
Basque (0-50)	34.38 (5.88)	46.04 (2.67)
Picture naming (0-65)		
Spanish	64.47 (0.92)	63.38 (1.62)
Basque	50.09 (9.63)	64.19 (1.47)
Interview (0-5)		
Spanish	5 (0)	5 (0)
Basque	4.14(0.47)	4.95(1.33)

Table 5: General proficiency assessment of the participants in the two groups

3.2 Stimuli

The visual stimuli included 50 pictures from 10 categories (5 per category: flute, horse, alarm-clock, bird, scissors, dog, camera, frog, door, and cow). Each of the 10 categories was represented by 5 different highly recognizable color images (.png extension, white background, 2000x2000 pixels): three photographs obtained from online image collections, one normed color drawing (Rossion & Pourtois, 2004), and one "cartoon" image (Lupyan & Thompson- Schill, 2012)[4]. Words were recorded by a Spanish-Basque female speaker and sounds were downloaded from online libraries. The mean number of letters was controlled among nouns (mean: 5.76 letters, SD 1.51; range: 4–9 letters). In addition, words have been checked for measures of concreteness (mean: 5.86, SD 0.58); imageability (mean: 6.11, SD 0.73) and familiarity (mean: 6.12.67, SD 0.53) (EsPal, Duchon et al., 2013). Half of the nouns referred to artifacts, and half to natural objects. The mean label/nonverbal sound length was 0.67 ± 0.05 s. In order to be sure that sounds and images were univocally identifiable, a group of Spanish-Basque bilingual speakers (N=20) who did

[4]*See the* Appendix *section to look at the pictures used.*

not take part in the experiment saw 100 images, and named them with the first noun that came up to their minds. Also, they listened to 50 different sounds (10 for each category). We only chose the images and the sounds whose name was univocally expressed by all the 20 participants.

3.3 Procedure

The ERP experiment was run in a soundproof electrically shielded chamber with a dim light. Participants were seated in a chair, about sixty centimeters in front of a computer screen. Stimuli were delivered with the PsychoPy software (Peirce, 2007).

Participants completed 200 trials of a cued-picture recognition task. On each trial, participants heard a word in Spanish, (e.g., *perro*, “dog”) or a nonverbal sound (e.g., a dog bark). Two seconds later (we selected this large time interval to make sure that ERPs came back to baseline prior to image presentation and indeed this was the case) the sound stimulus onset a picture appeared and participants responded “yes” or “no” by pressing two buttons on a keyboard to indicate whether the picture matched the auditory cue. In 50% of the trials the picture matched the auditory cue at the category level (perro “dog” → dog; [bark] → dog). In the remaining 50% (incongruent trials), the image that followed was from one of the other nine categories. The picture remained visible until a response was made. All trial parameters were fully randomized within participants. Overall, the experiment lasted 40 minutes on average.

3.4 Electrophysiological recording and data analysis

EEG was recorded from 27 electrodes placed in an elastic cap (Easycap, www.easycap.de): Fp1, Fp2, F7, F8, F3, F4, FC5, FC6, FC1, FC2, T7, T8, C3, C4, CP5, CP6, CP1, CP2, P3, P4, P7, P8, O1, O2, Fz, Cz, Pz. All sites were online referenced to the left mastoid (A1). Additional external electrodes were placed on mastoids (A1, A2) and around the eyes (VEOL, VEOR, HEOL, HEOR) in order to detect blinks and eye movements. A forehead electrode served as the ground. Data were amplified (Brain Amp DC) with a bandwidth of 0.01−100 Hz, at a sampling rate of 250 Hz. The impedance of the scalp electrodes was kept below 5 kΩ, while the eye electrodes impedance was below 10 kΩ. Collected recordings were off-line re-referenced to the average activity of the two mastoids. Artifacts exceeding 100 μV in amplitude were rejected. Raw data were visually inspected and artifacts such as muscular activity and ocular artifacts were marked for subsequent rejection.

For the analysis of the ERPs on the image, epochs of 1700 ms (from -200 ms to 1500ms) were obtained to select time windows of interest. For each condition, the average ERP waveforms were computed time-locked to the onset of the image. After baseline correction (−0.50 ms to 0 sec) epochs were averaged independently for each condition and subject.

In order to explore the presence of a P100 effect we selected a 80ms-long time interval of interest (from 90 to 170 ms) and entered in the software R the average ERP activity of the electrodes. We run a three-way ANOVA with Group (two levels: Spanish vs. Basque), Input (two levels: Spanish cue vs. Sound cue) and Electrode (only posterior electrodes were considered: O1, O2, P7, P8; similar to Boutonnet & Lupyan, 2015). Also, we took in consideration only the images that matched with the previous cue, in order to have a cleaner result. P-values were Greenhouse-Geisser corrected.

For the analysis of the N400, the statistics time-locked to the image focused on the time-interval classically related to lexical integration processing (300-500 ms, namely the N400). In this latter analysis as well we run a three-way ANOVA: Group, Input, Electrode) as an index of successful integration of the target image in the context.

4. Results

From the analysis on the P100 emerged a significant interaction between group and input [F (1,42) = 4,18 , p< 0.004, *ges*= 0.0003).

Pictures that were cued by labels elicited more positive P100 peak amplitudes than when the same pictures were cued by nonverbal sounds. Crucially, this effect was had no significant effect in the Basque (L1) speakers [F (1,21) = 1,60 , p>0.2, *ges*= 0.002).

There was a strong difference between groups on the ERP elicited by labels (*Spanish-Basque label: M= 2.39, SD=1.35; Basque-Spanish label: M= 1.84, SD= 1.92)*. On the

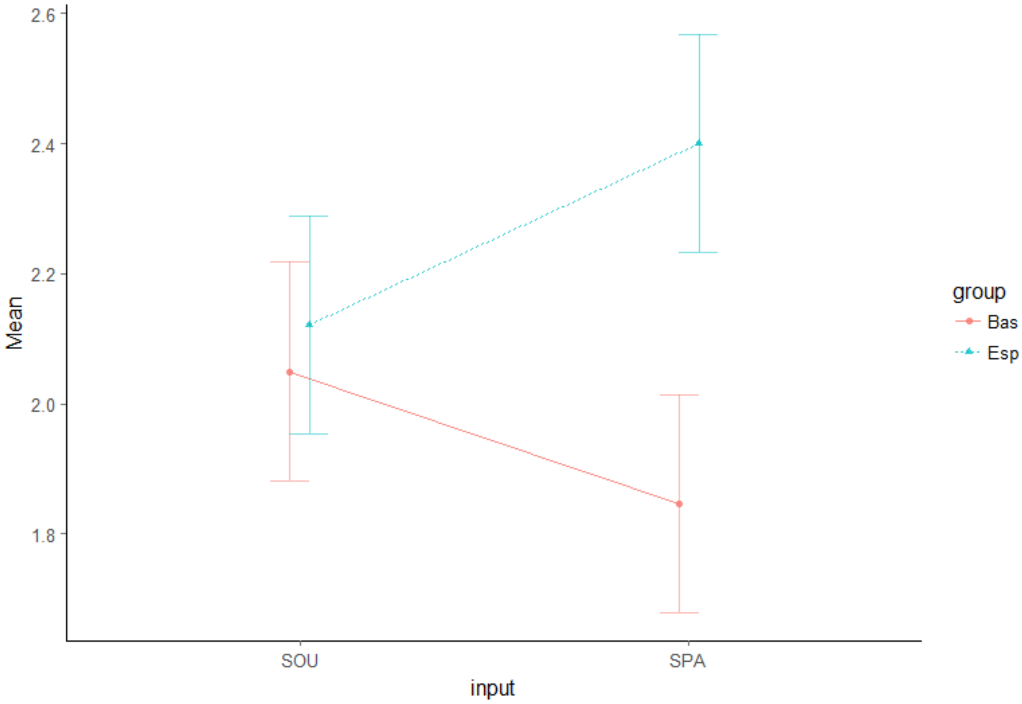


Table 6: Between Group interaction

other hand, there was no difference on the values elicited by non-verbal sounds (*Spanish-Basque non-verbal cue M= 2.12, SD= 1.25; Basque-Spanish non-verbal cue M= 2.04, SD=1.85)* (see Table 6). The analysis on the N400 component showed an expected result. The mismatch condition elicited a strong significant effect, in comparison with the match condition (the match effect was [F (1,46) = 79,81, $p < 0.001$, *ges*= 0.06]. Importantly, there was no group effect [F (1,46) = 0.54, $p > 0.46$, *ges*= 0.006], and no input effect [F (1,46) = 3.39 , $p > 0.07$, *ges*= 0.002]) meaning there was no difference between the signal elicited by non-verbal cued images and word-cued images in neither of the groups.

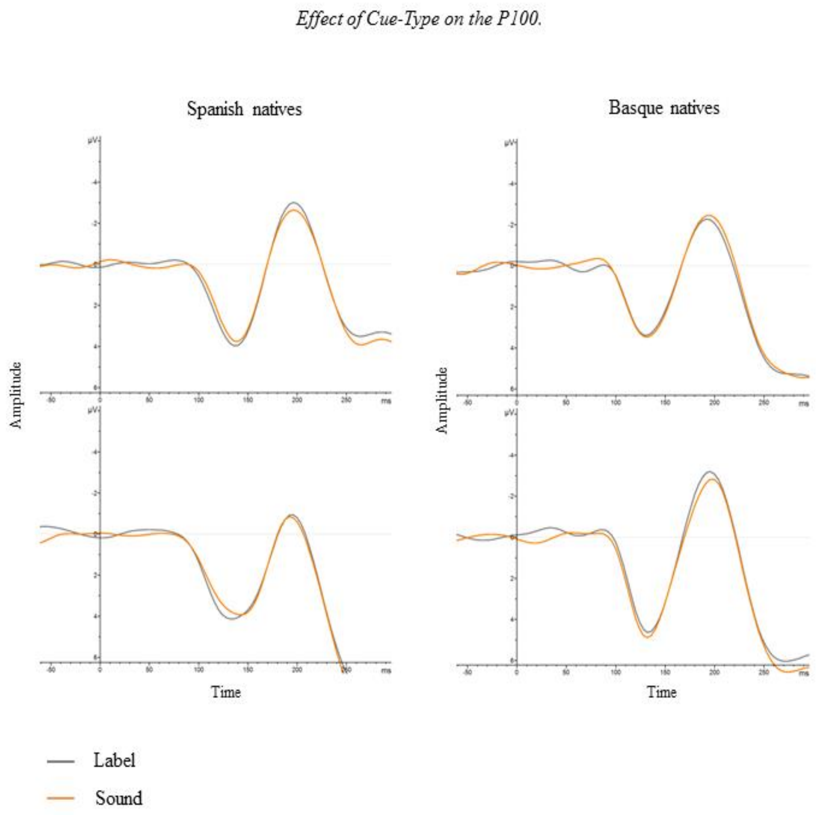


Figure 7: P100 ERPs for the two groups across the posterior electrodes (linear derivations of P7 and O1 for the left, and P8 and O2 for the right). In grey are ERPs elicited by label-cued pictures, and in orange are ERPs elicited by sound-cued pictures.

Effect of Congruence on the N400.

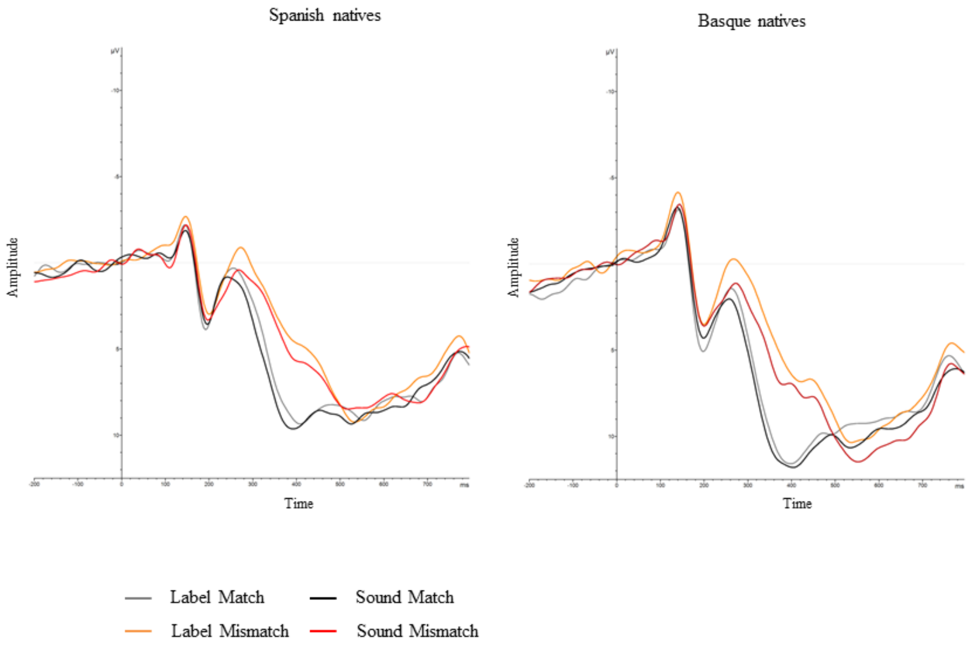


Figure 8: N400 ERPs for Spanish natives and Basque natives, across the two levels of cue type and congruence (linear derivations of FC1, FC2, C3, CZ, C4, CP1, CP2). In grey and black are matching trials. In red and orange are mismatching trials. Grey and orange lines are label-cued trials, red and black trials are sound-cued trials.

5. Discussion

According to previous studies (Boutonnet & Lupyan, 2015, inter alia), language does not only influence "high" semantic levels of cognition. Verbal labels can provide top-down signals to the visual system, changing how subsequently incoming (bottom-up) information is processed. In particular it has been found that people are better at recognizing images when they are cued by words than by equally familiar sounds.

In the present study we analyzed whether very early L2 speakers are able to activate fast visual predictions like natives do. A group of Spanish-Basque and Basque-Spanish bilinguals (with the L2 learnt at 3 y.o.) went through a word-picture matching task. They heard a Spanish word, or a sound of the same word's category. An image appeared after the auditory stimulus. Participants had to indicate whether the image matched or not with the stimulus they heard, while ERPs time-locked to the onset of the image were recorded.

We expected the Spanish natives to show an advantage in the processing of word-cued images in comparison with sounds-cued images, in line with Boutonnet & Lupyan (2015).

This is exactly what emerged from the analysis on the ERP data. Words elicited a more positive P100 effect on the images in comparison with the effect elicited by non-verbal sounds. This suggests that in bilinguals, after hearing a word in the L1, the neural processes responsible for generating the P100 were already sensitive to the object category.

On the other hand, the expectations for the Basque natives had two possible outcomes. We hypothesize they could display the same ERP results found on the Spanish natives, as a consequence of their high proficiency in Spanish, and show the same P100 effect on images elicited by verbal labels. Alternatively, their very high proficiency could represent a disadvantage, so they could perform not as well as Spanish and not show any early perceptual effect.

This second option is indeed the outcome emerged from the results. Basque natives did not display any P100 effect on the images cued by verbal labels: they do not activate any visual prediction about the input they perceive. In our opinion, Basque natives' knowledge of the Spanish lexicon was way too strong to negatively influence performance in this task. It is very unlikely that this outcome was due to the fact that the verbal-labels participants heard were in Spanish (L2). It is very likely that the cause of this outcome is Basque natives' high proficiency indeed, but for a different reason. Participants probably suffered a strong interference from their L1, the Basque language. At the very early visual level, the contrast between two highly activated languages did not permit to activate visual predictions fast enough. In contrast with the experiments presented in the previous chapters, here Basque natives do not show any advantage from their language interaction.

Basque natives did not show any prediction effect. Language interaction does not necessarily imply advantages in terms of language prediction, as emerged from the two previous studies. Instead, in the present experiment, the more balanced proficiency level between languages determined an interference which did not allow to the Basque natives to visually predict and caused a disadvantage. But why is that so?

Speculation about why two labels similarly activated do not lead to perceptual facilitation is still an open question. Even if different languages' lexicons are composed of very different word forms, these words describe the same world and thus shared conceptual structures. Still, acquiring native-like word knowledge of an L2 is not simply to match the known lexical items of the first language to those of the L2. This is because languages differ in the way that words or names are mapped onto objects (see Malt and Majid, 2013 for review).

According to Zinszer et al. (2011), proficient bilingual speakers develop patterns of naming that differ from the patterns of monolingual speakers of either language, suggesting that the bilinguals' lexical representations reflect the convergence of two languages toward each other. From their connectionist point of view, a way that languages influence each other is via the associative connection weights holding between the word meaning and the word form. When a new L2 word form is learned as a translation equivalent of an L1 word, the network will set weights to match those of the L1 word. The L2 word will be activated by the same features as the L1 word, resulting in non- native L2 patterns of production. Over time, these weights may settle into a pattern that does not exactly match that of the L2 but reflects the convergence of L1 and L2. In such an account, the meanings associated with word forms are different for the bilingual compared to those of monolingual speakers of the two languages.

Another possible form of influence can come from the associative connections of weights between two languages on the word form level. Starreveld et al. (2014) suggested that observed bilingual word use may not reflect convergence in underlying representations in the two languages. Instead, it may reflect connections between two sets of entirely monolingual-like representations such that cross-activation of representations of both languages results in one language having an influence on the output in the other. Thus, when a new L2 word form is learned as a translation equivalent of an L1 word, the word forms are linked, and activation of a word in one language will create activation to the associated word in the other language.

Both the explanations could apply to the results of our experiment: it could be that for balanced bilingual speakers translations across two languages lead to activation of slightly different semantic representations, and consequently not to unique perceptual expectation for the picture.

Finally, the analysis on the later time window, displayed an expected outcome. Both Spanish natives and Basque natives showed a similar N400 effect on the images that did not match with the previous stimuli, meaning the groups are not different in the semantic processing. Importantly, there was no difference between verbal and non-verbal cued mismatches, supporting the contention that the two cue types are equivalently informative at a semantic level in the context of this task.

6. Conclusion

We provided evidence that native speakers are faster to recognize an image after hearing a verbal label than after hearing a sound. Words are able to set priors that alter how a subsequent image is visually processed within 100 ms of stimulus onset. In other words, the target picture is processed in light of predictions triggered by verbal labels, but not by sounds. This result has not been found in highly proficient non-native speakers. While there is no different in the semantic processing between natives and non-natives, very fast visual prediction mechanisms are impaired in non-native speakers. This outcome is probably due to the interference between their L1 and their L2.

Chapter 5

General discussion

1. Lexical Prediction in bilingual speakers

Language processing strongly relies on predictive processing mechanisms (Bar, 2007; Clark, 2013; Federmeier, 2007; Friston, 2010; Levy, 2008; Pickering & Garrod, 2013; but see Jackendoff, 2002; Morris, 2006 for a different point of view). Predictive processes are a valuable resource for the speakers' processing system: they contribute to the ease and speed with which language comprehension incrementally builds upon contextual constraints and internal world knowledge (Federmeier, 2007; Levy, 2008; Pickering & Garrod, 2013). Some authors consider prediction so fundamental that it has been suggested that learning to speak in infants arises directly from learning to predict (Mani &Huettig, 2012).

As we could see in the first chapter, top-down operations not only facilitate the incorporation of new inputs into the preceding structure after they have been observed (*integration*), but also pre-activate upcoming inputs before they are even perceived (*prediction*; e.g., Altmann & Mirkovic, 2009; Jaeger & Snider, 2013; Kuperberg & Jaeger, 2016; Levy, 2008; MacDonald, 2013).

Anticipatory behaviour in the native language has been attested by a great number of studies.

Experiments using the visual word paradigm have shown that when given a constraining context, participants tend to move their eyes towards objects or images before they are explicitly revealed by the input (Altmann & Kamide, 1999; Kamide, Altmann & Haywood, 2003). Research using Event-Related Potentials (ERPs) manipulated context sentences in order to measure what happens in the brain right before participants perceive an unexpected lexical item (in comparison with an expected one), and found replicable electrophysiological correlates of lexical prediction, namely a more negative N400 effect (DeLong, Urbach & Kutas, 2005; Wicha, Moreno & Kutas, 2003). Also, it has been found that oscillatory EEG activity around 4 Hz reflects the degree to which a certain item can be predicted (Molinaro, Barraza, & Carreiras, 2013).

When they listen or read, native speakers are able to predict the phonological (DeLong et al., 2005), orthographic (Hawelka et al., 2015), syntactic (Levy, 2008; Dikker et al., 2010) and semantic (Altmann & Kamide, 1999; Kutas & Hillyard, 1984) representation of the input coming next. In the case of mismatch between the predicted and the actual input, language users take advantage of the *prediction error* to adjust future predictions, with the result of speeding up processing, enhancing communication and boosting learning (Chang, Dell, & Bock, 2006; Federmeier, 2007; Pickering & Garrod, 2013).

However, even if predictive sentence processing in the first language has been widely studied, a comprehensive picture of this phenomenon is still far from being complete.

Much is known about what native speakers are able to predict, but much less is known about when and how they do it. Language processing is extremely fast, therefore anticipation processes (and their relative time course) are very hard to grasp. In this research project we investigated the mechanisms supporting language prediction, along with their operational features. In our opinion, research on second language anticipation mechanisms was a valid approach to this issue and it was fundamental for comprehending the general mechanisms underlying language prediction in general. Research on prediction in bilingual speakers has provided clear evidence that low proficiency second language speakers, who do not rely on life-long experience, have not developed robust prediction processes as native speakers do (for a review, Kaan, 2014). When processing a message in an L2, most speakers are not fast enough to keep pace with the speed with which a sentence unfolds, and their comprehension is often impaired (Hahne, 2001; Hahne & Friederici, 2001). Initial research on bilingual prediction has shown that non-native speakers differ from natives in terms of quantity and quality of exposure to the L2: this leads to reduced proficiency and impaired prediction mechanisms (Kaan, 2014). Nonetheless, the efficiency of the prediction mechanism would directly depend on the proficiency in each language, therefore it seems reasonable to assume that when L2 proficiency is high enough, L2 prediction mechanism are very much like the ones used in the L1.

The present research provided a step further in the study of bilingual prediction. Until now, most experiments have focused on the presence/absence of prediction effects in late/low proficiency bilinguals. We assumed that prediction is possible in L2 as far as proficiency levels are high enough. This work explored the anticipation processes that are at play during L2 lexical pre-activation when L2 proficiency is very high (thus comparable with L1 prediction processes). Three main ERP experiments have been run. Through the analysis of bilingual speakers, our studies not only aimed at finding the differences between L1 and L2 in anticipation processing, but also had the goal of explaining these differences within a more general and unified theory of the prediction mechanisms underlying language processing.

In order to get started with our investigation, we took into consideration two main experiments from the limited literature that focused on second language prediction. The debate generated by the results of these two central studies gave us ground to base our research on.

In a first experiment, Martin et al. (2013) investigated lexical prediction comparing a group of late but high proficient Spanish-English bilinguals and a group of English monolinguals. They used an ERP paradigm similar to DeLong et al. (2005) to study the differences between L1 and L2 in the way semantic processing is modulated by lexical pre-activation. Participants read English highly constraining sentences like *He was very tired so he sat on…* By the time they read the sentence, subjects had already predicted the exact noun that would have come next. Nonetheless, sentences could be followed by either an expected or an unexpected (but correct) noun phrase (noun preceded by its article) such as *a chair* or *an armchair*. Crucially, the experiment utilized the phonological property of English according to which the indefinite article "*a*" changes to "*an*" if the noun that follows it starts with vowel, and took advantage of this rule to detect prediction.

In monolinguals, the ERPs recorded on the unexpected article (preceding unexpected nouns) elicited more negative N400 amplitudes than the expected article, in line with DeLong et al. (2005). In contrast, highly proficient bilinguals did not show any prediction effect in the L2. Based on this result, the authors concluded that L2 users do not predict to the same extent native speakers do.

This seemed a strong result but did not consider neither the language typological differences nor the cross-linguistic similarities.

In order to shed light on this matter, a follow up study (Foucart et al., 2014) compared Spanish monolinguals, Spanish-Catalan early bilinguals and French-Spanish late bilinguals. The experiment followed Wicha et al. (2004): participants read sentences in Spanish where the context was manipulated in a way that the critical noun and its preceding article could be expected or unexpected. Critically, the gender of the expected noun was the opposite of the unexpected ones, and their corresponding articles agreed with them (e.g., *The pirate had the secret map, but he never found the* [masc] *treasure* [masc] / *the* [fem] *cave* [fem]).

ERPs recorded on the unexpected article elicited a more negative N400 effect than the expected articles. Interestingly, the same N400 effect was found in all three groups of participants, regardless of the L1 they spoke, their L2 proficiency or their age of L2 acquisition.

In contrast with the previous experiment, in this study the conclusion was that bilinguals can predict upcoming words in the same way monolinguals do.

The main difference between the two experiments was the linguistic feature manipulated to detect prediction mechanisms. While in the first experiment the phonotactic agreement feature ("*a*", "*an*") exists in English but not in Spanish, the gender agreement feature between article and noun is present in Spanish, French and Catalan.

This final distinction opened up several central issues. In particular, it was unclear whether typological language differences have an impact on L2 prediction processes, or whether bilinguals were able to predict the features of the L2 even though those features are not part of their L1 system. More generally, we did not know if a native-like prediction capacity could be attained with a high proficiency level, or if the L1 characteristics could have an irrecoverable impact on L2 anticipation mechanisms.

2. Experiment 1

With the aim to solve these matters we run our first study. We tested Basque (L1)-Spanish (L2) and Spanish (L1)-Basque (L2) speakers; all of them were highly proficient in both languages but they were different in terms of age of acquisition, with the L2 acquired at the age of three years old.

As in the previous experiments (Foucart et al., 2014; Wicha et al., 2004), subjects read sentences (in Spanish) in which a constraining context was followed by an expected or an unexpected noun phrase, where the latter was formed by a gender marked article and a gender marked noun. Here again, expected and unexpected nouns differed in gender, so that it was possible to observe anticipation effects on the article. In comparison with the previous studies, however, there was a significant difference, as we also considered the hierarchical level at which the gender information is available for the predicted noun.

We took advantage of the typological difference between Spanish and Basque morphology. The Basque language does not assign grammatical gender to the nouns, and it is based on post-nominal suffixes (Laka, 1996; Rijk, 2008). On the other hand, to every

Spanish noun is assigned grammatical gender, but this feature is irregular in one third of the cases: most of the nouns are gender-transparent (ending in –a for feminine , and in -o for masculine), but the remaining ones are gender-opaque (ending in consonant, or in a different vowel). According to a theory by Gollan & Frost (2001), in order to extract gender language users can follow or the lexical route (based on the noun features stored in the mental lexicon), or the formal route (based on the formal properties of the noun). As gender of Spanish opaque nouns cannot be retrieved from the lexical route, speakers have to rely on the formal one, while the reliable noun-endings of the transparent nouns permit to extract the gender though both the formal and the lexical route (Caffarra et al., 2014).

Crucially, our first study did not only use the gender difference between expected and unexpected nouns to detect prediction, but we also considered the transparency feature. Both groups of participants read highly constraining sentences that could end with opaque or transparent nouns like *Acabo de salir de la casa y no recuerdo si he cerrado la puerta/el armario"* ("I just got out of the house , and I don't remember whether I closed the [fem] door [fem, transp] / the [masc] wardrobe [masc, transp]") or *Prefiero que el te esté muy dulce, puedes pasarme el azúcar / la miel por favor?* ("I prefer my tea very sweet, would you pass me the [masc] sugar [masc, opaque] / the [fem] honey [fem, opaque] ?"). ERPs were recorded both on the determiner onset and on the noun onset.

The experiment results were highly informative. First of all, both groups showed a strong N400 effect on the unexpected determiners as opposed to the expected ones. The effect was not connected to the transparency of the predicted noun. This result solved the main question emerged from the debate between Martin et al. (2013) and Foucart et al. (2014), proving that early bilinguals can predict like monolinguals even using a feature which is absent in their L1 (namely gender).

A second crucial outcome emerged from the results. Only Basque natives displayed an N200 like effect on the unexpected determiners preceding transparent nouns (but not on those preceding the opaque nouns).

To interpret this early effect, a language processing hypothesis was designed. Spanish natives rely on the lexical route to predict the gender of the nouns; as a consequence they do not take into account the transparency feature because they extract the gender from the mental lexicon and not from formal features. On the contrary, Basque natives are driven by the post-nominal suffix analysis they are used to run in their L1, and apply the same process to the L2. Therefore, we think that they depend more on the formal analysis for the prediction of the gender. This analysis is reflected in the N200 for the determiners preceding transparent words, but there is no trace of it on the ERPs related to determiners preceding the opaque nouns. It is very likely that they run the default sub-lexical analysis for all the nouns, and they succeed in the case of transparent noun, but they fail with the opaque nouns because they have no formal cues to base their formal analysis on. Still, they are very highly proficient in Spanish, so they switch to the lexical route for gender prediction.

If our hypothesis is correct this first study's results have got significant implications.

To start with, we know that the difference between the groups was mainly represented by the language to which participants were exposed during their first three years of life. Henceforth, the native language experience must have been the element that influenced the way participants predict gender. The linguistic regularities determined by the environment during infancy created the dependency on different levels of linguistic representations during prediction. This is definitely a step further in the research on bilingual prediction, because it goes beyond paradigms that until now focused on the presence or absence of L2 prediction depending on the proficiency, and proposes a more graded process made of stages. In fact, our hypothesis implies the existence of hierarchical levels of prediction, directly represented by earlier and later ERP components (and determined by the native language exposure): an earlier prediction level of visual-analysis, reflected by the N200 effect, and a later lexical-related effect represented by the N400. In addition, the study provides a first electrophysiological evidence that language prediction does not only deal with "high-level" linguistic representations, but that top-down information goes from internal representation to sensory regions, influencing perception. We will better talk about the relationship between prediction and perception in the discussion regarding the third experiment.

3. Experiment 2

The second experiment had the goal to find further evidence of the existence of hierarchical levels of prediction in language. We used two groups of bilinguals with the same characteristics of the previous study: participants were Spanish (L1) - Basque (L2) speakers and Basque (L1) - Spanish (L2) speakers, who acquired the L2 at the age of three years old. In the previous study, lexical prediction was triggered by constraining sentences, and the effects were recorded on the article preceding the predicted noun. A crucial difference in the second study was that we decided not to use sentence-context to create expectations, but we used a word marching task (WMT) and a picture matching task (PMT). This way we eliminated the possibility that sentences triggered prediction of the determiner (for its syntactic role) more than prediction of the content word following it. In a way, we created a more direct and simple way to detect lexical prediction. Participants read a Spanish noun or saw a picture on a screen. An auditory NP followed the stimuli, and participants had to indicate whether it matched or not with what they read or saw. We manipulated gender and transparency: mismatching nouns were different in gender compared to the presented stimuli; also, nouns could be transparent or opaque. In the NP, the determiner always agreed with the noun, but it was or congruent or incongruent in respect to the stimulus. ERPs were recorded on the article and on the noun.

We based our logic on composite language processing models. The idea was that, when participants read a word, the word-form features of that word get activated at a sub-lexical level. This activation should be reliably visible 200 ms after the word onset. Only later (about 400 ms) the lexical information of that word gets activated at a higher level (Grainger & Holcomb, 2009; Grainger & Jacobs, 1996; McClelland & Rumelhart, 1981). On the other hand, when participants see a picture, the lexical and semantic information connected to that image gets activated, and only after this activation the word-form can be derived (Indefrey & Levelt, 2004; Strijkers & Costa, 2016).

This is what happens in lexical processing, but we wanted to know what happens in lexical prediction, so we tracked the time course of the mismatching effect time-locked to the determiner preceding the predicted word. Crucially, the activation depended on the format of the predictor, that could be a written word, or a picture.

Let us put it simpler. Participants perceive a stimuli (word or image), and they know they have to perform a matching task, so they wait for the auditory stimuli corresponding to the perceived item to come. At this stage they pre-activate not only the word but also the gender features of that word (as clearly emerged from the previous study). The idea is that in the WMT the word activates sequentially the sub-lexical information first, and then the lexical information. As a consequence, when they perceive a determiner disagreeing with the stimuli, the mismatching gender information should be perceived at the sub-lexical stage, it should be electrophysiologically visible earlier in time, and it could be considered as a prediction effect. Instead, the PMT processing is different because the image activates sequentially the lexical information before the sub-lexical one, so the prediction effect on the mismatching article should be visible earlier than in the WMT, but still lexically-related.

In addition, this experiment not only manipulated gender but transparency as well. Opaque nouns do not provide the participants with any formal cue about their gender, so the only way to extract it is through the lexical route. Both Spanish and Basque natives should display the same pre-activation timing (earlier in the PMT than in the WMT, as the lexical information is available earlier in the former).

On the contrary, we expect transparent nouns to determine differences between groups. If Spanish natives rely on the lexical information for gender prediction, transparent nouns prediction should elicit a lexical effect on the unexpected articles both in the WMT and in the PMT. If Basque natives rely on sub-lexical information for gender prediction, in the WMT we should see an earlier effect on the unexpected determiner, in comparison with the Spanish natives. As in the PMT the lexical information is accessed before the sub-lexical, we expect Basque natives to display a delayed effect compared to the Spanish natives, because they would run an additional sub-lexical analysis. The results of this study confirmed our expectations.

In the WMT, both Spanish and Basque natives showed the same N400 effect on unexpected articles preceding opaque words. Determiners preceding transparent words displayed instead a different pattern. In both groups there was a negativity connected to a lexical effect at ~390 ms. Crucially, only in Basque natives the unexpected determiner preceding transparent words also elicited an earlier effect at ~240 ms, to confirm that this group relies on sub-lexical information for gender prediction of transparent words.

In the PMT, the unexpected determiners coming before the opaque words showed a similar effect at ~250 ms for both groups, in confirmation of the fact that opaque words can only be predicted through lexical information. ERPs recorded on determiners preceding transparent words showed an effect at ~430 ms in the Spanish group, and later effect at ~550 ms in the Basque group. The difference is very likely due to the fact that Basque natives performed additional sub-lexical analysis after the lexical one, resulting in later timing.

This study confirmed that language prediction is tuned to the characteristics of the first language. In order to predict Spanish words, Basque natives continue to apply the regularities of their L1, even though they have been exposed to Spanish all lifelong. Spanish natives, instead show a strong mapping of the grammatical gender stored in the mental lexicon.

More importantly, this experiment achieved the goal of providing further evidence for the existence of hierarchical levels in language prediction. Gender can be predicted on at least two levels, with an early stage in which language users can rely on the formal information of the upcoming item, and another stage operating later, where gender can be extracted from the mental lexicon.

These results replicated the previous study, and did it by using a completely different task, as context was not created by sentences but single items. In addition, the present predictees were audio and not visual (as they were in the previous experiment), giving us poof that prediction is also modality independent.

4. Experiment 3

As emerged in the first two experiments, Basque natives showed a processing advantage in comparison with Spanish natives. Basque natives could rely both on the sub-lexical and on the lexical information for the prediction of the gender of Spanish nouns, proving to be better predictors. We assumed this phenomenon to be related to the typological characteristics of the Basque language.

In order to test whether Basque natives are better predictors in general, or whether the L1 actually interferes with the processing and the prediction of the L2, we run a control study.

We analyzed two groups of participants with the same characteristics of the first two studies. Spanish (L1) - Basque (L2) and Basque (L1) - Spanish (L2) first exposed to the L2 at the age of three were our subjects. As the previous results were very likely due to the capacity to process morphology faster, we eliminated this last factor and analyzed prediction by a different angle. We studied the relationship between language prediction and visual perception.

According to behavioral and electrophysiological studies, native speakers are faster and better at recognizing a picture of a dog after hearing the word "dog" than after hearing a bark (Boutonnet & Lupyan, 2015). In a predictive account this means that verbal labels, when perceived, provide the visual system with a set of priors that bias the process of incoming stimuli. In contrast, non verbal labels do not seem to do it. Language prediction influences visual perception. Boutonnet & Lupyan (2015) concluded that words denote categories, abstracting the peculiarities of that word, while non-verbal sounds are necessarily linked to the specifics of that noun therefore less categorical ("dog" denotes all dogs, barking reminds the sound of a particular dog).

We replicated Boutonnet & Lupyan (2015), but we used out two groups of bilingual speakers in order to understand whether bilinguals visual system would be advantaged by verbal labels in the L1, or whether L1 would interfere with L2 prediction. This experiment gave us the possibility to eliminate the "morphology" factor, and analyze lexical prediction at a rawer, faster level of processing. Participants heard a word in Spanish (e.g., *perro*, "dog") or a non-verbal cue (e.g., dog bark). After the auditory stimuli, a picture appeared on a screen. The picture could match or not with the auditory cue. Participants had to indicate whether the picture matched with what they heard. ERPs time-locked to the onset of the image were measured.

In Spanish natives words elicited a more positive P100 effect on the images in comparison with the effect elicited by non-verbal sounds. Basque natives did not display any P100 effect on the images cued by verbal labels: they did not activate any visual prediction about the input they perceived. This outcome was probably due to their high proficiency indeed. Participants suffered a strong interference from their L1, the Basque language. At the very early visual level, the contrast between two highly activated languages did not permit to activate visual predictions fast enough. In contrast with the experiments presented in the previous chapters, here Basque natives did not show any advantage from their language interaction.

5. Interpretations and limits of the study

Altogether, the results of the three experiments that form this thesis contribute to elaborate a model describing the way lexical prediction might work.

In order to generate predictions, we provided the participants with top-down pieces of information. We gradually reduced the amount of information, and changed its format, so we could be able to analyze prediction mechanisms from progressively closer angles. In the first experiment the context was represented by highly constraining sentences that could lead to the anticipation of only one item. In the second experiment, WMT's context was represented by a single word that provided enough linguistic information to perform the task. In addition, PMT's context was represented by a single image, with the purpose of testing non-linguistically triggered linguistic items. Finally, in the third experiment, prediction was generated by simple word labels that were free of any morphosyntactic information, as opposed to sound labels.

The different top-down pieces of information generated predictions at three different levels of activation, that are hierarchically represented in terms of neural timing. At a very early stage (the "sub-lexical" level), the word-form analysis of the upcoming word is performed. The operation takes place around 200 ms after the onset of an eliciting context-item. Visual attention mechanisms are dedicated to the formal analysis of an upcoming word, in order to better extract its morphological features. At a second prediction stage, a later fully lexical analysis is reflected by N400-like effects. The concept related to the upcoming word, along with its morphological features, is predicted and extracted here. Finally, around 100 ms after the onset of a verbal auditory stimulus, a visual prediction of an upcoming image is generated. At this level, that we could call "visual" stage, the predictions produced by the context can alter the perception of the bottom-up input: the effect is present only when anticipation is generated by a linguistic item and not by a sound.

Clearly, language prediction is an extremely fluid and plastic process, and its hierarchy is very far from being such a schematic phenomenon because it is influenced and shaped by many independent factors.

First of all comes the language factor. We could be able to detect the different prediction levels only through the comparison of our two groups of Spanish-Basque and Basque-Spanish speakers. Their native language experience, and the characteristics of their L1, as long as the interaction between their two languages, determined their predictive performance. Both groups showed the presence of a lexical level of prediction analysis, but only the Basque natives were able to run a prediction at a sub-lexical level, thanks to the their L1 typology. Prediction at the "visual" stage, instead, was only present in the Spanish natives performing in the L1, and not in the Basque natives, because of the interference between the languages they had to cope with.

Secondly, the prediction levels are differently triggered by the quality of distinct top-down information, so that prediction effects can smoothly move along the hierarchy and adapt to it. For example, in the PMT the prediction of opaque words was performed through lexical analysis by both groups. Importantly, the timing of this effect emerged much earlier than the one displayed in the WMT (at 250 ms in the former, as opposed to the classical 400 ms in the latter). The earlier effect could only be due to the task, in fact the top-down information provided (namely the image) forced an access to the lexical level before the sub-lexical one.

Third, we found evidence that the hierarchical levels can communicate among them and interface each other. A good example was present again in the PMT. The prediction of transparent words determined a lexical effect on the articles emerging at 340 ms in Spanish natives, but the effect appeared only at 460 ms in Basque natives. The time delay in Basque natives is very likely due to additional processing from a lower lexical level. In fact they could not avoid using more support from the sub-lexical information level: the result is an effect happening ~130 ms later than that of the Spanish natives.

The plasticity of language prediction is in no doubt. Nonetheless, the first years of linguistic exposure play a fundamental role in shaping prediction mechanisms in a way that is no longer changeable, even at very high levels of L2 proficiency. Basque natives predicting Spanish words exhibited sub-lexical analyses as a consequence of their L1: they could not avoid running this analysis. In some cases this could represent an advantage, as for the prediction of transparent nouns, but it could also determine a disadvantageous interference at the very early stages of prediction, like for the verbal labels as opposed to the sound cues. We can state that language prediction is plastic indeed, but some mechanisms learned in the first years of life can hardly be changed during life.

To sum up, the characteristics of the languages spoken, the quality of the top-down information, the language proficiency, and the native experience are the main factors influencing prediction. In order to rapidly adapt to these factors, prediction mechanisms get differently shaped in an extremely plastic way. This process can lead to advantages and disadvantages in terms of prediction efficiency.

This research has got a few limits. The first is that, for all the experiments, we analyzed bilinguals who spoke the same two languages, namely Basque and Spanish. As the differences in processing emerged for Basque natives were due to the typological characteristics of the Basque language, it is very likely that we would not find similar results with different bilinguals. Future research should definitely aim at studying bilinguals speaking languages with different typological features, and testing the existence of prediction hierarchies among them.

Our experiment tested morphological prediction of the gender feature, and showed the strategy bilingual brain can use to predict grammatical gender. However, we can safely suppose that prediction mechanisms are able to predict several other morphosyntactic traits of the upcoming items. By using diverse speakers, with different languages, future studies should point at exploring which features can be anticipated, when in the time hierarchy they would be processed, and how prediction plasticity would move in order to compute them.

As we saw, the top-down information available to the speakers determines and modifies the prediction time-course. In a future research different kind of contexts eliciting prediction should be conceived, in order to better understand the way top-down information influence the perception of bottom-up stimuli.

Finally, we only tested very early bilinguals, and found evidence of prediction mechanisms that are unchangeable, even with very high proficiency level. It would be very interesting to study bilinguals with different ages of L2 acquisition, and different levels of proficiency in order to shed light on how the interplay between these two factors can influence prediction mechanisms in bilinguals. Furthermore, longitudinal studies need to be conducted to examine how the anticipatory use of information changes as an L2 speaker becomes more proficient and how individuals differ.

6. Conclusion

Overall, this thesis gives a great contribution to the research on bilingual prediction and on the mechanisms beneath language prediction in general. Taking advantage of the typological distance between Spanish and Basque in early balanced bilinguals, we add important pieces of evidence to the puzzle on how bilingualism shapes language prediction.

First of all, we provided evidence that early bilinguals are able to predict L2 nouns, along with their grammatical features. Crucially, we also found that they are able to predict the features of the L2 nouns even if those features are totally absent in their L1 system.

Secondly, we demonstrated that the native language experience (even for only the first three years of life) largely shapes prediction mechanisms, so that bilinguals reading in their second language rely on the distributional regularities that are highly relevant in their first language in order to predict.

More importantly, we showed that the individual linguistic experience hierarchically modulates the format of the predicted representation. In fact we provided critical evidence that linguistic prediction is essentially hierarchical, with different levels of operation differently represented in terms of neural timing that are able to communicate at an interface stage.

In other words, this research shows that predictions can rapidly adapt to new contexts, selecting all the available cues in order to build an internal representation of the new environment: early linguistic experience models the way in which different cues are weighted to obtain optimal predictions. The whole process is extremely flexible, with neural timing constantly ready to adapt to the environmental demands and to mediate between native experience settings and new language feature requirements.

Finally we saw that the plasticity of language prediction does not always represent an advantage. We found that bilingual speakers are not able to make visual predictions like native speakers do. Language interaction can work performing some tasks, as we saw in the case of Spanish morphology for Basque natives. Nonetheless, language interference can represent a problem in the case of very fast prediction mechanisms like the visual ones. In this case the L1 can interfere with the L2, and impair prediction.

Acknowledgements

This work was partially supported by the Spanish Ministry of Economy and Competitiveness (grants PSI2012-32350 and PSI2015-65694-P to N.M.; grant PSI2014-54500-P to C.M. and S.C.; and grant PSI2015-65694-P to N.M.), by the Basque government (grant PSI2015-1-25 to C.M. and S.C.) and by the Ikerbasque Research Fellowships to N.M and C.M. The project was also partially supported by the award ‘‘Centro de Excelencia Severo Ochoa SEV-2015-0490”. Further support derived from the AThEME project funded by the European Commission 7th Framework Programme, the ERC-2011-ADG-295362 from the European Research Council.

Firstly, I would like to express my sincere gratitude to my tutor Prof. Maria Teresa Guasti, for believing in me during these hard years of study, for the continuous support during my Ph.D.research, for her extraordinary patience, and immense knowledge.

I will forever be thankful to my advisor Prof. Nicola Molinaro. His guidance helped me in all the time of research and writing of this thesis. He allowed me to grow as a research scientist and as a person: his advice have been precious and invaluable. His enthusiasm and love for science is contagious. I could not have imagined having a better mentor for my Ph.D study.

My sincere thanks also goes to Dr. Sendy Caffarra and Dr.Clara Martin, who provided me the opportunity to join their team, and work with them. Without their precious help and support it would have not been possible to conduct my research.

I would like to thank the “Proactive Group” for the brilliant comments, the feedback, and the suggestions during the creation and the interpretation of my experiments.

Also, I wish to express my gratitude to the BCBL lab-staff and the research-assistants who helped to recruit the participants, to collect the data and helped me to get results of better quality.

I would like to thank all the fantastic people and the great scientists that I met during these four years both in Milan and in San Sebastiàn: they have all helped and taught me immensely. Among them I had the luck to have true friends. I thank my friends for the life they gave me the opportunity to live. I am grateful of sharing with them joys, tears, laughs, fears, and love.

Finally, I especially thank my mom, dad, sister and granma. My hard-working parents have sacrificed their lives for my sister and myself and provided unconditional and immeasurable love and care. They have cherished with me every great moment and supported me whenever I needed it. I love them so much, and I would not have made it this far without them.

References

Abney, S. P. (1989). A computational model of human parsing [Special issue] . *Journal of Psycholinguistic Research* , *18* (1), 129 –144 .

Allison, T., Puce, A., Spencer D.D., McCarthy, G. (1999). Electrophysiological studies of human face perception. Potentials generated in occipitotemporal cortex by face and non-face stimuli. *Cerebral Cortex, 9*, 415– 430.

Altmann, G. T., & Kamide, Y. (1999). Incremental interpretation at verbs: Restricting the domain of subsequent reference. *Cognition*, 73(3), 247–264.

Altmann, G. T., & Mirkovic, J. (2009). Incrementality and prediction in human sentence processing. *Cognitive Science, 33(4)*, 583-609.

Altmann , G. T. M. , & Steedman , M. J. (1988). Interaction with context during human sentence processing . *Cognition* , *30*, 191–238 .

Arnal, L. H., Doelling, K. B., & Poeppel, D. (2015). Delta–beta coupled oscillations underlie temporal prediction accuracy. *Cerebral Cortex, 25*(9), 3077–3085.

Bar, M. (2007). The proactive brain: Using analogies and associations to generate predictions. *Trends in Cognitive Sciences, 11*(7), 280–289.

Bastos, A. M., Vezoli, J., Bosman, C. A., Schoffelen, J. M., Oostenveld, R., Dowdall, J. R., Fries, P. (2015). Visual areas exert feedforward and feedback influences through distinct frequency channels. *Neuron*, *85*(2), 390–401.

Boersma, P. (2001). Praat, a system for doing phonetics by computer. *Glot International 5:9/10*, 341-345.

Boutonnet, B., Dering, B., Vinas-Guasch, N., Thierry, G. (2013) Seeing objects through the language glass. *Journal of Cognitive Neuroscience 25*, 1702–1710.

Bloom, P. (2000). How children learn the meanings of words. Cambridge, MA: MIT.

Caffarra, S., & Barber, H. (2015). Does the ending matter? The role of gender-to-ending consistency in sentence reading. *Brain Research*, *1605*, 83-92.

Caffarra, S., Janssen, N., & Barber, H. (2014). Two sides of gender: ERP evidence for the presence of two routes during gender agreement processing. *Neuropsychologia, 63,* 124-134.

Caramazza, A. (1997). How many levels of processing are there in lexical access? *Cognitive Neuropsychology,14,* 177–208.

Carroll, J., & Bever , T . (1978). Th e perception of relations: A case study in the relation of knowledge and perception. In E. Carterette & M. Friedman (Eds.), *Handbook of perception: Vol . 7. Speech and language* . New York: Academic Press.

Chang, F., Dell, G. S., & Bock, K. (2006). Becoming syntactic. *Psychological Review, 113*(2), 234–272.

Clahsen, H., & Felser, C. (2006). Grammatical processing in language learners. *Applied Psycholinguistics, 27,* 3–42.

Clark, A. (2013). Whatever next? Predictive brains, situated agents, and the future of cognitive science. *Behavioral and Brain Sciences, 36*(3), 181–204.

Coltheart, M., Rastle, K., Perry, C., Langdon, R., Ziegler, J. (2001). DRC: A dual route cascaded model of visual word recognition and reading aloud. *Psychological Review, 108,* 204–256.

Daneman, M., & Carpenter , P. A. (1983). Individual differences in integrating information between and within sentences . *Journal of Experimental Psychology: Learning, Memory, and Cognition* , *9* , 561–584 .

Cooper, R. M. (1974). The control of eye fixation by the meaning of spoken language: A new methodology for the real-time investigation of speech perception, memory, and language processing. *Cognitive Psychology*, 6 (1), 84-107.

de Rijk, R. P. G., & de Coene, A. (2008). *Standard Basque: A Progressive Grammar*. Cambridge, MA: MIT Press.

Dell, G. S., & Chang, F. (2014). The P-chain: Relating sentence production and its disorders to comprehension and acquisition. *Philosophical Transactions of the Royal Society B: Biological Sciences, 369* (1634), 1–9.

DeLong, K., Urbach, T. P., & Kutas, M. (2005). Probabilistic word pre-activation during language comprehension inferred from electrical brain activity. *Nature Neuroscience, 8* (8), 1117–1121.

Dehaene, S. Duhamel, J., Hauser, M., and Rizzolatti, G. (2005) *From Monkey Brain to Human Brain.* A Fyssen Foundation Symposium. MIT Press.

Dell, G. S., & Brown, P. M. (1991). Mechanisms for listener-adaptation in language production: Limiting the role of the “model of the listener". In D. J. Napoli & J. A. Kegl (Eds.), *Bridges between psychology and linguistics: A Swarthmore Festschrift for Lila Gleitman* (Vol. 105, pp. 105–129): Psychology Press.

Dijkstra, T., Grainger, J., & van Heuven, W. J. B. (1999). Recognition of cognates and interlingual homographs: The neglected role of phonology. *Journal of Memory and Language*, *41*(4), 496–518.

Dikker, S., Rabagliati, H., Farmer, T. A., & Pylkkänen, L. (2010). Early occipital sensitivity to syntactic category is based on form typicality. *Psychological Science, 21(5),* 629-634.

Duchon, A., Perea, M., Sebastián-Gallés, N., Martí, A., Carreiras, M. (2013). EsPal: one- stop shopping for Spanish word properties. *Behavioral Research Methods, 45*(4), 1246-1258.

Dussias, P. E. (2003). Syntactic ambiguity resolution in L2 learners: Some effects of bilinguality on L1 and L2 processing strategies. *Studies in Second Language Acquisition, 25*, 529–557.

Dussias, P. E. (2004). Parsing a first language like a second: The erosion of L1 parsing strategies in Spanish-English bilinguals. *International Journal of Bilingualism, 3,* 355–371.

Dussias, P. E., & Pinar, P. (2010). Effects of reading span and plausibility in the reanalysis of wh- gaps by Chinese-English second language speakers. *Second Language Research, 26*(4), 443–472.

Edmiston P, Lupyan G (2013) Verbal and nonverbal cues activate concepts differently, at different times. Proceedings of the 35th Annual Conference of the Cognitive Science Society, Berlin, July–August.

Federmeier, K. D. (2007). Thinking ahead: The role and roots of prediction in language comprehension. *Psychophysiology, 44*, 491-505.

Federmeier, K. D., McLennan, D. B., De Ochoa, E., Kutas, M. (2002). The impact of semantic memory organization and sentence context information on spoken language processing by younger and older adults: an ERP study. *Psychophysiology, 39,* 133-146.

Ferreira, F., Foucart, A., & Engelhardt, P. E. (2013). Language processing in the visual world: Effects of preview, visual complexity, and prediction. *Journal of Memory and Language*, *69*(3), 165–182.

Firestone, C., Scholl, B.J. (2014) “Top-down” effects where none should be found: the El Greco fallacy in perception research. *Psychological Science, 25,* 38–46.

Fodor, Jerry A. (1983). *Modularity of Mind: An Essay on Faculty Psychology*. Cambridge, Mass.: MIT Press.

Foucart, A., Martin, C. D., Moreno, E. M., & Costa, A. (2014). Can bilinguals see it coming? Word anticipation in L2 sentence reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *40*(5), 1461–1469.

Forster, K. (1979). Levels of processing and the structure of the language processor. In W. E. Cooper & E. C. T. Walker (Eds.), *Sentence processing* (pp. 25-86). Hillsdale, NJ: Erlbaum.

Francken, J.C., Kok, P., Hagoort, P., de Lange, F.P. (2015) The behavioral and neural effects of language on motion perception. *Journal of Cognitive Neuroscience,* 27, 175–184.

Friston, K. (2005). A theory of cortical responses. *Philosophical Transactions of the Royal Society B: Biological Sciences, 360*(1456), 815–836.

Friston, K. J. (2010). The free-energy principle: a unified brain theory? *Nature Reviews Neuroscience, 11(2),* 127-138.

Gibson, E., & Pearlmutter, N. J. (2000). Distinguishing serial and parallel parsing. *Journal of Psycholinguistic Research, 29*(2), 231-240.

Gervain, J., & Meheler, J. (2010). Speech perception and language acquisition in the first year of life. *Annual Reviews of Psychology, 61*, 191–218.

Gervain, J., Sebastian-Galles, N., Diaz, B., Laka, I., Mazuka, R., Yamane, N., Nespor, M., & Mehler, J. (2013). Word frequency cues word order in adults: cross- linguistic evidence. *Frontiers in Psychology, 4,* 689.

Gleitman, L., Papafragou, A. (2005) *Language and thought*. In: Cambridge handbook of thinking and reasoning (Holyoak KJ, Morrison B, eds), pp 633–661. Cambridge: Cambridge University.

Grainger, J., & Holcomb, P. J. (2009). Watching the word go by: On the time-course of component processes in visual word recognition. *Language and Linguistic Compass, 3(1),* 128-156.

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

Gollan, T., & Frost, R. (2001). Two Routes to Grammatical Gender: Evidence from Hebrew. *Journal of Psycholinguistic Research, 30*, 627–651.

Gollan, T. H., Montoya, R. I., Cera, C., & Sandoval, T. C. (2008). More use almost always means a smaller frequency effect: Aging, bilingualism, and the weaker links hypothesis. *Journal of Memory and Language, 58* (7), 87–814.

Gollan, T. H., Slattery, T. J., Goldenberg, D., Van Assche, E., Duyck, W., & Rayner, K. (2011). Frequency drives lexical access in reading but not in speaking: The Frequency-Lag Hypothesis. *Journal of Experimental Psychology: General, 140* (2), 186–209.

Guthrie, D., & Buchwald, J.S. (1991). Significance testing of difference potentials. *Psychophysiology, 28*(2), 240–244.

Hahne, A. (2001). What's different in second-language processing? Evidence from event-related brain potentials. *Journal of Psycholinguistic Research, 30,* 251–266.

Hahne, A., & Friederici, A.D. (2001). Processing a second language: Late learners comprehension mechanisms as revealed by event-related brain potentials. *Bilingualism: Language and Cognition, 4(2),* 123-141.

Hale, J. (2001). *A probabilistic Earley parser as a psycholinguistic model*. Paper presented at the Proceedings of the North American Chapter of the Association for Computational Linguistics on Language technologies (NAACL '01).

Hale, J. (2003). The information conveyed by words in sentences. *Journal of Psycholinguistic Research, 32*(2), 101-123.

Hale, J. (2011). What a rational parser would do. *Cognitive Science*, *35*(3), 399-443.

Hawelka, S., Schuster, S., Gagl, B., & Hutzler, F. (2015). On forward inferences of fast and slow readers. An eye movement study. *Scientific reports, 5.*

Harris, J. W. (1991). The exponence of gender in Spanish. *Linguistic Inquiry, 22,* 27–62.

Hoeks, J.C.J., Stowe , L.A., & Doedens, G. (2004). Seeing words in context: The interaction of lexical and sentence level information during reading . *Cognitive Brain Research* , *19* (1), 59–73 .

Hopp, H. (2013). Grammatical gender in adult L2 acquisition: Relations between lexical and syntactic variability. *Second Language Research*, *29* (1), 33–56.

Indefrey P., Levelt W. J. M. (2004). The spatial and temporal signatures of word production components. *Cognition 92,* 101–144.

Ito, A., Martin, A.E., & Nieuwland, M.S. (in press). On prediction of form and meaning in a second language. *Journal of Experimental Psychology: Learning, Memory, and Cognition.*

Izura, C., Cuetos, F., & Brysbaert, M. (2014). Lextale-Esp: A test to rapidly and efficiently assess the Spanish vocabulary size. *Psicologica, 35*, 49-66.

Jackendoff, R. (2002). *Foundations of Language: Brain, Meaning, Grammar, Evolution*. Oxford: Oxford University Press. p. 477

Jaeger, T. F., & Snider, N. E. (2013). Alignment as a consequence of expectation adaptation: syntactic priming is affected by the prime's prediction error given both prior and recent experience. *Cognition, 127(1)*, 57-83.

Janssen, N., Hernández-Cabrera, J. A., van der Meij, M., & Barber, H. A. (2015). Tracking the Time Course of Competition During Word Production: Evidence for a Post-Retrieval Mechanism of Conflict Resolution. *Cerebral Cortex, 25*(9), 2960, 2969.

Jurafsky, D. (1996). A probabilistic model of lexical and syntactic access and disambiguation. *Cognitive Science*, *20*(2), 137-194.

Just, M.A., & Carpenter, P.A. (1980). A theory of reading: From eye fixations to comprehension . *Psychological Review*, *87*, 329–354 .

Kaan, E. (2014). Predictive sentence processing in L1 and L2. *Linguistic Approaches to Bilingualism*, *4*(2), 257-282.

Kamide, Y., Altmann, G. T., & Haywood, S. L. (2003). The time-course of prediction in incremental sentence processing: Evidence from anticipatory eye movements. *Journal of Memory and Language, 49*, 133–156.

King, W. K., & Kutas, M. (1998). Neural plasticity in the dynamics of human visual word recognition. *Neuroscience Letters, 244*, 61– 64.

Kintsch, W., & van Dijk, T.A. (1978). Towards a model of text comprehension and productions . *Psychological Review , 85,* 363–394.

Klemfuss, N., Prinzmetal, W., Ivry, R.B. (2012). How does language change perception: a cautionary note. *Frontiers of Psychology, 3,* 78.

Kok, P., de Lange, F.P. (2014). Shape perception simultaneously up- and downregulates neural activity in the primary visual cortex. *Current Biology*. *24,* 1531– 1535.

Kok, P., Failing, M.F., de Lange, F.P. (2014) Prior expectations evoke stimulus templates in the primary visual cortex. *Journal of Cognitive Neuroscience, 26,*1546 –1554.

Kotz, S.A. (2009). A critical review of ERP and fMRI evidence on L2 syntactic processing. *Brain and Language, 109,* 68–74.

Kuperberg, G. R., & Jaeger, T. F. (2016). What do we mean by prediction in language comprehension? *Language, Cognition and Neuroscience*, *31*(1), 32-59.

Kutas, M. , & Hillyard, S.A. (1980). Reading senseless sentences: Brain potentials reflect semantic incongruity . *Science , 207* (4427), 20 –205.

Kutas, M. , & Hillyard, S.A. (1983). Event-related brain potentials to grammatical errors and semantic anomalies . *Memory and Cognition , 11*, 539–550.

Kutas, M., & Hillyard, S. A. (1984). Brain potentials during reading reflect word expectancy and semantic association. *Nature, 307(5947),* 161-163.

Laka, I. (1996). *A Brief Grammar of Euskara, the Basque Language*. Euskal Herriko Unibertsitatea: Leioa-Donostia (Spain)

Lemhofer, K., & Broersma, M. (2012). Introducing LexTALE: A quick and valid Lexical Test for Advanced Learners of English. *Behavioral Research Methods, 44*(2), 325-343.

Levelt, W. J. M., Roelofs, A., & Meyer, A. S. (1999). A theory of lexical access in speech production. *Behavioral and Brain Sciences*, *22*, 1–38.

Levy, R. (2008). Expectation-based syntactic comprehension. *Cognition, 106* (3), 1126–1177.

Lewis, A. G., & Bastiaansen, M. (2015). A predictive coding framework for rapid neural dynamics during sentence-level language comprehension. *Cortex, 68*, 155-168.

Li, P., Dunham, Y., Carey, S. (2009) Of substance: the nature of language effects on entity construal. *Cognitive Psychology 58,* 487–524.

Luck, S.J. (2005). *An Introduction to the Event-Related Potential Technique*. MIT Press, Cambridge, MA.

Luck S.J., Hillyard S.A. (1994). Electrophysiological correlates of feature analysis during visual search. *Psychophysiology*, *31,* 291–308.

Luke, S.G., & Christianson, K. (2016). Limits on lexical prediction during reading. *Cognitive Psychology, 88,* 22-60.

Lupyan, G., Thompson-Schill, S.L. (2012). The evocative power of words: Activation of concepts by verbal and nonverbal means. *Journal of Experimental Psychology*. *141*, 170 – 186.

Lupyan, G. (2012a). Linguistically modulated perception and cognition: the label-feedback hypothesis. *Frontiers in Psychology, 3*,54.

Lupyan, G. (2012b). *What do words do? toward a theory of languageaugmented thought*. In: Psychology of learning and motivation (Ross BH, ed), pp 255–297. New York: Elsevier.

Lupyan, G., Bergen, B. (2015). *How language programs the mind. Topics in Cognitive Science*. New Frontiers in Language Evolution and Development. In press.

Lupyan, G., Clark, A. (2015). Words and the world: predictive coding and the language-perception-cognition interface. *Current Directions in Psychology*. In press.

Lupyan, G., Thompson-Schill, S.L. (2012). The evocative power of words: Activation of concepts by verbal and nonverbal means. *Journal of Experimental Psychology: General, 141*, 170 –186.

Lupyan, G., Rakison, D.H., McClelland, J.L. (2007) Language is not just for talking: redundant labels facilitate learning of novel categories. *Psychological Science*, *18*, 1077–1083.

MacDonald, M. (2013). How language production shapes language form and comprehension. *Frontiers in Psychology, 4*.

Malt B. C., Majid A. (2013). How thought is mapped into words. *Wiley Interdisciplinary Reviews: Cognitive Science* 4583–597.

Mani, N., & Huettig, F. (2012). Prediction during language processing is a piece of cake – but only for skilled producers. *Journal of Experimental Psychology: Human Perception and Performance, 38*(4), 843–847.

Mangun, G., Hillyard, S.A. (1991). Modulations of sensory-evoked brain potentials indicate changes in perceptual processing during visual-spatialpriming. *Journal of Experimental Psychology: Human Perception and Performance, 17,* 1057–1074.

Maris, E., & Oostenveld, R. (2007). Nonparametric statistical testing of EEG- and MEG-data. *Journal of Neuroscience Methods*, *164*(1), 177-190.

Marslen-Wilson, W.D. (1987). Functional parallelism in spoken word recognition. *Cognition 25(1-2)*, 71-102.

Marslen-Wilson, W. D., & Tyler , L. K. (1980). The temporal structure of spoken language understanding. *Cognition , 8* , 1–71 .

Martin, C., Thierry, G., Kuipers, J.-R., Boutonnet, B., Foucart, A., & Costa, A. (2013). Bilinguals reading in their second language do not predict upcoming words as native readers do. *Journal of Memory and Language*, *69*(4), 574–588.

McClelland, J. L., & Rumelhart, D. E. (1981). An interactive activation model of context effects in letter perception: part I: an account of basic findings. *Psychological Review, 88*, 375.

Meyer A.M., Federmeier K.D. Event-related potentials reveal the effects of aging on meaning selection and revision. *Psychophysiology* in press.

Michalareas, G., Vezoli, J., van Pelt, S., Schoffelen, J. M., Kennedy, H., & Fries, P. (2016). Alpha-Beta and Gamma Rhythms Subserve Feedback and Feedforward Influences among Human Visual Cortical Areas. *Neuron, 89*(2), 384-397.

Mitchell , D.C. , & Green , D.W. (1978). The effects of context and content on immediate processing in reading . *Quarterly Journal of Experimental Psychology , 30*, 609 – 636.

Mitsugi, S., & MacWhinney, B. (2015). The use of case marking for predictive processing in second language Japanese. *Bilingualism: Language and Cognition, 19*(1), 19–35.

Molinaro, N., Barber, H.A., Pérez, A., Parkkonen, L., & Carreiras, M. (2013a). Left fronto-temporal dynamics during agreement processing: Evidence for feature-specific computations. *Neuroimage*, *78*, 339-352.

Molinaro, N., Barraza, P., & Carreiras, M. (2013b). Long-range neural synchronization supports fast and efficient reading: EEG correlates of processing expected words in sentences. *Neuroimage*, *72*, 120-132.

Molinaro, N., Carreiras, C., & Duñabeitia, J.A. (2012). Semantic combinatorial processing of non-anomalous expressions. *Neuroimage*, *59*(4), 3488-3501.

Molinaro, N., Conrad, M., Barber, H.A., & Carreiras, M. (2010). On the functional nature of the N400: Contrasting effects related to visual word recognition and contextual semantic integration. *Cognitive Neuroscience, 1*(1), 1-7.

Molinaro, N., Monsalve, I., & Lizarazu, M. (2016). Is there a common oscillatory brain mechanism for producing and predicting language? *Language, Cognition and Neuroscience, 31*(1), 145-158.

Molinaro, N., Vespignani, F., Canal, P., Fonda, S., & Cacciari, C. (2008). Cloze-probability does not only affect N400 amplitude: The case of complex prepositions. *Psychophysiology*, *45*(6), 1008-1012.

Molnar, M., Lallier, M., & Carreiras, M., (2014). The Amount of Language Exposure Determines Nonlinguistic Tone Grouping Biases in Infants From a Bilingual Environment. *Language Learning, 64*(s2), 45-64.

Morris , R. K. (2006). Lexical processing and sentence context effects . In M. J. Traxler & M. A. Gernsbacher (Eds.), *Handbook of psycholinguistics* (2nd ed., pp. 377–402).

Oostenveld, R., Fries, P., Maris, E., & Schoffelen, J.M. (2011) FieldTrip: Open Source Software for Advanced Analysis of MEG, EEG, and Invasive Electrophysiological Data. *Computational Intelligence and Neuroscience, 2011*, 156869.

Osterhout, L., Bersick, M., & McKinnon, R. (1997). Brain potentials elicited by words: Word length and frequency predict the latency of an early negativity. *Biological Psychology, 46*, 143–168.

Pickering, M. J., & Garrod, S. (2007). Do people use language production to make predictions during comprehension? *Trends in Cognitive Sciences, 11,* 105–110.

Pickering, M. J., & Garrod, S. (2013). An integrated theory of language production and comprehension. *Behavioral and Brain Sciences, 36*(4), 329–347.

Peirce, JW (2007) PsychoPy-Psychophysics software in Python. *Journal of Neuroscience Methods, 162(1-2)*, 8-13.

Pritchett , B. L. (1991). Head position and parsing ambiguity [Special issue] . *Journal of Psycholinguistic Research ,20* (3), 251 – 270 .

R Core Team (2015). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. Vienna, Austria.

Rao, R. P., & Ballard, D. H. (1999). Predictive coding in the visual cortex: a functional interpretation of some extra-classical receptive-field effects. *Nature Neuroscience, 2(1),* 79-87.

Real Academia Española. (2009). Nueva gramática de la lengua española. Asociaciòn de academias de la lengua Española.

Rescorla, R. A., & Wagner, A. R. (1972). A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement. In A. H. Black & W. F. Prokasy (eds.), *Classical conditioning II: current research and theory* (pp. 64-99) New York: Appleton-Century-Crofts.

Rijk, R.P.G. (2008). *Standard Basque, a Progressive Grammar*. Cambridge MA: MIT Press.

Roelofs, A., (1992). A spreading-activation theory of lemma retrieval in speaking. *Cognition 42,* 107-142

Rossi, S., Gugler, M. F., Friederici, A. D., & Hahne, A. (2006). The impact of proficiency on syntactic second-language processing of German and Italian: Evidence from event-related potentials. *Journal of Cognitive Neuroscience, 18*, 2030–2048.

Rossion B, Pourtois G (2004) Revisiting Snodgrass and Vanderwart's object pictorial set: the role of surface detail in basic-level object recognition. *Perception, 33*,217–236.

Schultz, W., Dayan, P., & Montague, P. R. (1997). A neural substrate of prediction and reward. *Science, 275*(5306), 1593-1599.

Sohoglu, E., & Davis, M. H. (2016). Perceptual learning of degraded speech by minimizing prediction error. *PNAS.*

Spivey, M. J., & Marian, V. (1999). Cross talk between native and second languages: Partial activation of an irrelevant lexicon. *Psychological Science*, *10* (3), 281–284.

Stanovich, K. E. & West, R. F. (1979) Mechanisms of sentence context effects: Automatic activation and conscious attention. *Memory and Cognition 7,* 77– 85.

Stanovich, K. E., & West, R. F. (1983). On priming by a sentence context . *Journal of Experimental Psychology: General , 112* (1), 1–36 .

Starreveld, P. A., De Groot, A. M. B., Rossmark, M. M., & Van Hell, J. G. (2014). Parallel language activation during word processing in bilinguals: Evidence from word production in sentence context. *Bilingualism: Language and Cognition, 17*, 258-276.

Staub, A., & Clifton, C., Jr. (2006). Syntactic prediction in language comprehension: Evidence from either…or . *Journal of Experimental Psychology: Learning, Memory, and Cognition* , *32* (2), 425– 436.

Steedman , M . (1989). Grammar, interpretation, and processing from the lexicon . In W. Marslen- Wilson (Ed.), *Lexical representation and process*. Cambridge, MA: MIT Press.

Steinhauer, K., White, E. & Drury, J.E. (2009). Temporal dynamics of late second language acquisition: Evidence from event-related brain potentials. *Second Language Research*, 25 (1), 13-41

Strijkers K, Costa A (2016). On words and brains: linking psycholinguistics with neural dynamics in speech production. *Language, Cognition and Neuroscience, 31(4)*, 524-35.

Taylor , W. L. (1953). "Cloze procedure": A new tool for measuring readability . *Journalism Quarterly* , *30* , 415 – 433 .

Thierry, G., Athanasopoulos, P., Wiggett, A., Dering, B., Kuipers, J.R. (2009) Unconscious effects of language-specific terminology on preattentive color perception. *Proceedings of the National Academy of Sciences, 106*, 4567– 4570.

Tokowicz, N., & MacWhinney, B. (2005). Implicit and explicit measuresnof sensitivity to violations in second language grammar: An event related potential investigation. *Studies in Second Language Acquisition, 27,* 173–204.

Traxler , M. , Bybee , M. , & Pickering , M. (1997). Influence of connectives on language comprehension: Eye-tracking evidence for incremental interpretation . *Quarterly Journal of Experimental Psychology A* , *50* (3), 481–497 .

Traxler, M. J., Pickering, M. J., & Clifton, C. (1998). Adjunct attachment is not a form of lexical ambiguity resolution. *Journal of Memory and Language, 39*(4), 558- 592.

Traxler, M. J. (2014). Trends in syntactic parsing: anticipation, Bayesian estimation, and good-enough parsing. *Trends in Cognitive Sciences, 18*(11), 605-611.

Tyler , L. K., & Marslen-Wilson, W. D. (1977). The on-line eff ects of semantic context on syntactic processing . *Journal of Verbal Learning and Verbal Behavior* , *16* , 683–692 .

van Berkum, J. J. A., Brown, C. M., & Hagoort, P. (1999). Early referential context effects in sentence processing: Evidence from event-related brain potentials . *Journal of Memory and Language* , *41* (2), 147–182

Van Hell, J., & Dijkstra, T. (2002). Foreign language knowledge can influence native language performance in exclusively native contexts. *Psychonomic Bulletin & Review, 9*(4), 780–789.

Van Petten, C., & Kutas, M. (1990). Interactions between sentence context and word frequency in event-related brain potentials. *Memory and Cognition* , *18* (4), 380–393.

Wang, X. J. (2010). Neurophysiological and computational principles of cortical rhythms in cognition. *Physiological Reviews, 90*(3), 1195–1268.

Weber, A., & Cutler, A. (2004). Lexical competition in non-native spoken-word recognition. *Journal of Memory & Language*, *50*, 1–25.

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*, 231–256.

Wicha, N. Y. Y., Moreno, E. M., & Kutas, M. (2003). Expecting gender: An event related brain potential study on the role of grammatical gender in comprehending a line drawing within a written sentence in Spanish. *Cortex, 39,* 483–508.

Wicha, N. Y. Y., Moreno, E. M., & Kutas, M. (2004). Anticipating words and their gender: An event-related brain potential study of semantic integration, gender expectancy, and gender agreement in spanish sentence reading. *Journal of Cognitive Neuroscience*, *16*(7), 1272– 1288.

Zinszer B. D., Malt B. C., Ameel E., Li P. (2011). Bilingual categorization behavior: internally or externally emergent? *Talk Presented at the 8th International Symposium of Bilingualism*, Oslo, Norway

Zinszer B. D., Malt B. C., Ameel E., Li P. (2012). Predictors of native-like categorization in Chinese learners of English. *Talk Presented at Second Language Research Forum* Pittsburgh, PA.

Appendix

Experiment 1

Expected NP

- A algunos presos no se les ve la piel, por todos los tatuajes que llevan en *la cara* dibujados.
- A esa hora ya no había transporte urbano, así que tuvo que llamar a *un taxi* sin remedio
- A mi abuela, le gustaba conocer la información así que cada día veía el informativo en *la tele* en la cocina.
- A mi padre le gusta estar al tanto de la información, por eso, cada día lee *la revista* sobre economía.
- Acabo de salir de la casa y no recuerdo si he cerrado *la puerta* cuando me he ido.
- Acudió a una pitonisa para que le leyera las cartas y ver que le deparaba *la fortuna* en el amor
- Al cabo de los días le llegó una multa por haber superado *la velocidad* permitida
- Al final de la película, el príncipe se acerca a la princesa y le da *un beso* con afecto.
- Al ver a su padre, preparó la mejilla porque esperaba recibir *una torta* por lo que había hecho
- Antes de cocinar el pollo, hay que precalentar *el horno* durante diez minutos.
- Antes de entrar al piso tuvo que quedar con el propietario para firmar *la escritura* ante notario.
- Antes la mili era obligatoria, pero ahora es voluntario alistarse en *el ejército* nacional.
- Cada invierno se hace una campaña para vacunar a la gente mayor contra *el virus* de la gripe.
- Cerca del hospital, escuchó el sonido de la sirena acercándose, así que apartó el coche para que pasase *el vehículo* rápidamente.
- Como quería escribir sus secretos, fue a la librería a comprarse *un diario* personal.
- Cuando dio la conferencia se quedo callado sin que le saliera ni *un suspiro* de la boca
- Cuando era joven tocaba la batería en *una banda* famosa
- Cuando era pequeño, al irse a la cama siempre le contaban *una historia* para dormir
- Cuando era sólo un niño mi abuelo plantó en este bosque *una flor* que ha crecido mucho.
- Cuando estaban en el salón viendo la televisión, siempre discutían para ver quién controlaba *el mando* en casa
- Cuando estés deshorientado, mira la brújula, porque siempre muestra *la dirección* y el camino.

- Cuando fue a pagar se metió la mano en el bolsillo y se dio cuenta de que le habían robado *la cartera* en el bar.
- Cuando ganaron la final, el capitán del equipo levantó *la copa* de Europa
- Cuando llevas muchos años con tu pareja, es muy difícil cortar *la relación* de repente
- Cuando un crucero se está hundiendo, primero abandonan la nave los pasajeros y en último lugar la tripulaciónsin duda
- De los cuatro elementos, el agua es el que mejor apaga *la llama* sin ninguna duda.
- Desde la terraza del apartamento de la playa se podía ver *la catedral* y el mar
- Desde que la dejó su novio, ya no cree en *el amor* verdadero.
- Desde su sofá, mi abuela espía lo que pasa fuera mirando por *el tejado* todo el día.
- Después de haber pisado una araña venenosa tuvo una irritación en *el pie* para una semana.
- Después de terminar el bachiller debes pensar si quieres estudiar en *la universidad* más cercana.
- Después del accidente no es capaz de recordar nada porque ha perdido *la memoria* totalmente.
- El abogado no estaba de acuerdo con la sentencia que había dictado el juez imminentemente
- El asesinato por arma blanca se produjo en la cocina, el arma homicida es presuntamente *un cuchillo* que estaba en la encimera
- El camión había quedado totalmente destrozado después de sufrir *una colisión* en la calle
- El desacuerdo diplomático se había prolongado tanto que los dos países armados estaban a punto de empezar *un conflicto* terrible
- El empresario despidió al trabajador porque se le terminó *el contrato* temporal
- El hinduismo, como el Islam, es *una religión* muy importante
- El hombre había pasado toda la noche emborrachándose en *el bar* del barrio
- El ladrón se quedó paralizado cuando vio al policía apuntándole con *el dedo* directamente
- El monumento de hierro más conocido de París es *la torre* Eiffel
- El rey llevaba en la cabeza *una corona* antigua.
- El símbolo del catolicismo es *el pez* en muchas iglesias.
- En el antiguo Egipto para honrar a cada Faraón que moría, construían *un ataúd de oro*.
- En el desayuno presentó las tostadas en un plato de porcelana y el café en *un jarro* de plástico.
- En el mapa que tenían los piratas la cruz indicaba donde estaba *la perla* mágica.
- En las grandes ciudades, el transporte más rápido es *la bicicleta* seguramente.
- Era alérgico a la lactosa, por eso le sientan mal la leche, el yogurt y *la crema pastelera*.

- Era de noche y tenía hambre, así que le preguntó a su madre cuánto faltaba para que se sirviera *el pollo* finalmente
- Era enero, hacía mucho frío y estaba solo en la montaña, pero creía que podía encontrarla, sólo tenía que seguir sus huellas en *el césped* fresco
- Era muy pequeño y no podía esquiar, así que le regalaron un trineo para que jugase en *la nieve* como todos.
- Es más fácil practicar un deporte minoritario cuando eres socio de *una asociación* subvencionada.
- Está tan delgada que se le nota la columna vertebral si le pasas la mano por *la espalda* desnuda
- Estaba cuidando sus plantas y la única que aguantaba sin agua era *el cactus* grande del rincón
- Estaba harto de tanto romanticismo, y es que en todas las canciones aparece *la amistad* eterna
- Este invierno ha habido muchas tormentas, el tejado de la cabaña se lo llevó *la lluvia* torrencial.
- Fue a ver un chalet que tenía un árbol con un columpio en *el jardín* de atrás.
- Había llenado hasta arriba la baniera para darse *un baño* relajante
- Había obtenido una puntuación casi perfecta, sólo había cometido *una imprecision* en la ortografia
- Había pasado tanto tiempo en un pueblo pequeño que, ahora que vivía en Nueva York, se alegraba de estar en *un lugar* tan grande
- Habían quedado en el cine para ver *una película* de acción
- Hace horas que no me meto nada en el estómago. Voy a la cocina porque no se me pasa *el hambre* constante
- Iba a la iglesia cada domingo, pero desde que murió su hijo ha dejado ir porque ha perdido *la fe* en Dios
- Iban por la autopista y para cruzar la montaña tuvieron que pasar por *una bifurcación* complicada.
- Intentó colgar el cuadro en *el salón* pero no le salió bien.
- La base de la mayoría de los platos asiáticos es *la carne* de ternera
- La comida estaba preparada así que la madre ordenó al hijo que pusiera *la mesa* lista
- La gallina se colocó sobre la piedra, dándole calor como si fuese *un huevo* suyo
- La niña no pesa nada, es ligera como *una pluma* de verdad.

- La ropa está sucia, ponla en *el suelo* por favor.
- La única manera de pasar al otro lado del río es cruzar *el puente* de piedra
- Le da miedo la oscuridad, nunca quiere quedarse sola en casa cuando llega *el final* del dia.
- Le encantaba el arte así que se dispuso a colgar en el centro de la pared *una obra* que había comprado.
- Le encantaba la musica clásica y tenía los dedos muy largos, así que decidió aprender a tocar *el piano* en la orquesta
- Le gustan mucho los instrumentos a cuerda y la música rock, así que toca *la guitarra* en un grupo.
- Le llamaron los directores de varios bancos después de que le tocase mucho dinero en *el sorteo* de navidad
- Lleva una falda fucsia, llama muchísimo *el interés* de los diseñadores.
- Lo que más odia de madrugar es el momento en que suena *la señal* a las 6
- Los astronautas flotan en el espacio porque allí no existe *la gravedad* de la tierra
- Los astronautas pasaron muchos meses en el espacio, sin poder abandonar *la nave* espacial para nada
- Los niños hacen castillos de arena en *el patio* durante el recreo.
- Los pescadores fueron recibidos por sus mujeres cuando el barco entró en *la bahía* de la Concha.
- Me quedo sin batería. Tengo el cargador conectado al móvil pero todavía me falta encontrar *un enchufe* adaptable.
- Ni la fruta, ni la verdura, ni el pescado, lo único que le gustaba era *el pastel* de chocolate.
- No había visto la señal y le pusieron una multa por cometer *un incidente* grave
- No he podido leer el libro y mañana es el examen, puedes hacerme *un resumen*, por favor?
- No paraba de llorar, pero él decía que sólo se le había metido algo en *la pupila* por el viento
- No podía entrar en el despacho porque había olvidado *el maletín* que contenía las llaves
- No puedo seguir escribiendo en el dictado, porque se le había terminado *el bolígrafo* negro
- No sabe cocinar, entonces para celebrar San Valentín la invitó a cenar en *un restaurante* de un hotel
- No todo es físico, una persona completa debe cultivar el cuerpo y *la mente* por igual
- Nunca mete un gol, no sabe jugar bien, pero le gusta mucho *el fútbol* profesional.
- Nunca sé dónde llevar mi móvil y mi cartera, tengo que comprarme *un bolso* que combine con todo.
- Para cortar la carne se necesita *una tabla* y un cuchillo.

- Para ir a Barcelona no puedes coger ni el avión, ni el autobús, sólo tienes que coger *el tren* desde aquí
- Para matar a un vampiro hay que clavarle una estaca en *el corazón* con fuerza
- Para pedirle matrimonio se arrodilló ante ella y le dio *un anillo* dorado
- Para ventilar la habitación mientras hace la cama, abre durante unos minutos *la ventana* pequeña
- Pepe era un chico muy activo, jugaba al futbol, al baloncesto, a la pelota, cada día practicaba *una clase* de deporte
- Por haber robadovarias cosas en una tienda, pasará un año presoen *la cárcel* de Basauri
- Por precaución, siempre que cojas la moto debes ponerte *la chaqueta* de piel
- Prefiero que el té esté muy dulce, puedes pasarme *el azúcar* de caña?
- Sabíamos que estaba casada porque llevaba un anillo en *el dedo*, que nunca se quitaba.
- Se despertó sudando y temblando, había tenido *una pesadilla* terrible.
- Se estaba terminando el tiempo para patinar sobre hielo, iban a cerrar *la pista* en diez minutos
- Si buscas a Ana, está preparando la comida, supongo que estará en *el mercado* comprando cosas.
- Si no se afeita con regularidad le crece mucho *la barba* y el bigote.
- Soy vegetariano, pero este jamón tiene muy buena pinta, voy a hacer *un disparate* esta vez.
- Su mujer se fue despues de 10 años de matrimonio. Eso le rompió *la ilusión* con la familia
- Subió la jaula del pájaro a un mueble temiendo que intentaría cazarlo *el gato* de los vecinos
- Temía que si se portaba mal, al morir acabaría en *el infierno* seguramente
- Tenía mucha sed, pero por fortuna se topó en mitad de la calle con *una fuente* abierta
- Tenía muchos cuadros, y quería poner un póster tambien, pero no le quedaba ningún hueco en *la pared* de cartón-yeso
- Tenía solo un 5% de probabilidades de sobrevivir, pero el todavía era optimista, como dice el refrán lo último que se pierde es *el ánimo* de vivir
- Tiene un resfriado fuerte. El médico le dio una receta de antibióticos y un jarabepara *el dolor* constante
- Todavía tenían un par de minutos de descuento antes de que el árbitro pitara *la terminación*del partido
- Todo quedó a oscuras porque se había ido *el sol* de pronto.
- Tuvieron que llamar a la grúa para que les remolcara *la roulotte* de repente
- Uno de los ingredientes básicos de un bonbón es *el chocolate* según la receta

- Ya tenía puesto el traje, pero tuvo que llamar a su mujer para que le ayudara a atarse *la corbata* roja

Unexpected NP

- A algunos presos no se les ve la piel, por todos los tatuajes que llevan en *el cuerpo* entero.
- A esa hora ya no había transporte urbano, así que tuvo que llamar a *una asistente* para que fuera a buscarla
- A mi abuela, le gustaba estar informada así que cada día veía el informativo en *el transistor* de la cocina.
- A mi padre le gusta estar al tanto de la información, por eso, cada día lee *el periódico* sobre economía
- Acabo de salir de la casa y no recuerdo si he cerrado *el armario* cuando me he ido.
- Acudió a una pitonisa para que le leyera las cartas y ver qué le deparaba *el futuro* en el amor
- Al cabo de los días le llegó una multa por haber superado *ellímite* de velocidad
- Al final de la película, el príncipe se acerca a la princesa y le da *una corona* con afecto.
- Al ver a su padre, preparó la mejilla porque esperaba recibir *un beso* de bienvenida.
- Antes de cocinar el pollo, hay que precalentar *la placa* eléctricadurante diez minutos.
- Antes de entrar al piso tuvo que quedar con el propietario para firmar *el contrato* ante notario
- Antes la mili era obligatoria, pero ahora es voluntario alistarse en *la marina* nacional.
- Cada invierno se hace una campaña para vacunar a la gente mayor contra *la gripe* común.
- Cerca del hospital, escuchó el sonido de la sirena acercándose, así que apartó el coche para que pasase *la ambulancia* rápidamente.
- Como quería escribir sus secretos, fue a la librería a comprarse *una pluma* y una libreta
- Cuando dio la conferencia se quedo callado, no le salía ni *unapalabra* de la boca
- Cuando era joven tocaba la batería en *un grupo* famoso
- Cuando era pequeño, al irse a la cama siempre le contaban *un cuento* para dormir
- Cuando era sólo un niño mi abuelo plantó en este bosque *un árbol* que ha crecido mucho.
- Cuando estaban en el salón viendo la televisión, siempre discutían para ver quién controlaba *la silla* más grande
- Cuando estés deshorientado, mira la brújula, porque siempre muestra *el norte* y el camino.

- Cuando fue a pagar se metió la mano en el bolsillo y se dio cuenta de que le habían robado *el diario* en el bar.
- Cuando ganaron la final, el capitán del equipo levantó *el puño con fuerza*
- Cuando llevas muchos años con tu pareja, es muy difícil cortar *el cordón* de repente
- Cuando un crucero se está hundiendo, primero abandonan la nave los pasajeros y en último lugar el capitán sin duda
- De los cuatro elementos, el agua es el que mejor apaga *el fuego* sin ninguna duda.
- Desde la terraza del apartamento de la playa se podía ver *el mar* y los surfistas
- Desde que la dejó su novio, ya no cree en *la felicidad* verdadera.
- Desde su sofá, mi abuela espía lo que pasa fuera mirando por *la ventana* todo el día.
- Después de haber pisado una araña venenosa tuvo una irritación en *la piel* durante una semana.
- Después de terminar el bachiller debes pensar si quieres estudiar en e*l país* más cercano.
- Después del accidente no puede recordar nada porque ha perdido *el conocimiento* totalmente.
- El abogado no estaba de acuerdo con la sentencia que había dictado *la corte* imminentemente
- El asesinato por arma blanca se produjo en la cocina, el arma homicida es presuntamente *una espada* que estaba en la encimera
- El camión había quedado totalmente destrozado después de sufrir *un accidente* en la calle
- El desacuerdo diplomático se había prolongado tanto que los dos países armados estaban a punto de empezar *una guerra* terrible .
- El empresario despidió al trabajador porque se le terminó *la baja* por paternidad
- El hinduismo, como el Islam, es *el pilar más* importante de algunas sociedades
- El hombre había pasado toda la noche emborrachándose en *la sede* del partido
- El ladrón se quedó paralizado cuando vio al policía apuntándole con *la pistola* directamente
- El monumento de hierro más conocido de París es *el portal* del museo Orsay.
- El rey llevaba en la cabeza *un sombrero* antiguo.
- El símbolo del catolicismo es *la cruz* en muchas iglesias.
- En el antiguo Egipto para honrar a cada Faraón que moría, construían *una pirámide* en el desierto.
- En el desayuno presentó las tostadas en un plato de porcelana y el café en *una taza* de barro
- En el mapa que tenían los piratas la cruz indicaba donde estaba *el tesoro* secreto.
- En las grandes ciudades, el transporte más rápido es *el metro* seguramente.

- Era alérgico a la lactosa, por eso le sientan mal la leche, el yogurt y *el queso* de vaca.
- Era de noche y tenía hambre, así que le preguntó a su madre cuánto faltaba para que se sirviera *la cena* finalmente
- Era enero, hacía mucho frío y estaba solo en la montaña, pero creía que podía encontrarla, sólo tenía que seguir sus huellas en *la nieve* fresca
- Era muy pequeño y no podía esquiar, así que le regalaron un trineo para que jugase en *el monte* como todos.
- Es más fácil practicar un deporte minoritario cuando eres socio de *un club* subvencionado.
- Está tan delgada que se le nota la columna vertebral si le pasas la mano por *el lomo* desnudo
- Estaba cuidando sus plantas y la única que aguantaba sin agua era *la flor* grande del rincón
- Estaba harto de tanto romanticismo, y es que en todas las canciones aparece *el amor* eterno
- Este invierno ha habido muchas tormentas, el tejado de la cabaña se lo llevó *el viento* una mañana.
- Fue a ver un chalet que tenía un árbol con un columpio en *la imagen* del catalogo de la agencia
- Había llenado hasta arriba la bañera para darse *una hora* de descanso
- Había obtenido una puntuación casi perfecta, sólo había cometido *un error* en la ortografia
- Había pasado tanto tiempo en un pueblo pequeño que, ahora que vivía en Nueva York, se alegraba de estar en *la ciudad* mas grande de todas
- Habían quedado en el cine para ver *un estreno* de temporada
- Hace horas que no me meto nada en el estómago. Voy a la cocina ya porque no se me pasa *la sed* constante
- Iba a la iglesia cada domingo, pero desde que murió su hijo ha dejado de ir porque ha perdido *el interés* en Dios
- Iban por la autopista y para cruzar la montana tuvieron que pasar por *un túnel* muy largo.
- Intentó colgar el cuadro en *la pared* pero no le salió bien.
- La base de la mayoría de los platos asiáticos es *el arroz* blanco
- La comida estaba preparada así que la madre ordenó al hijo que pusiera *el juego a un lado*
- La gallina se colocó sobre la piedra, dándole calor como si fuese *una cría* suya
- La niña no pesa nada, es ligera como *un pájaro* de verdad.
- La ropa está sucia, ponla en *la lavadora* por favor.
- La única manera de pasar al otro lado del río es cruzar *la calle* de piedra
- Le da miedo la oscuridad, nunca quiere quedarse sola en casa cuando llega *la noche* silenciosa

- Le encantaba el arte así que se dispuso a colgar en el centro de la pared *un cuadro* que había comprado.
- Le encantaba la musica clásica y tenía los dedos muy largos, así que decidió aprender a tocar *la viola* en la orquesta
- Le gustan mucho los instrumentos de cuerda y la música rock, entonces toca *el bajo* muy bien.
- Le llamaron los directores de varios bancos después de que le tocase mucho dinero en *la lotería* de navidad
- Lleva una falda fucsia, llama muchísimo *la atención* de todos.
- Lo que más odia de madrugar es el momento en que suena *el despertador* a las 6
- Los astronautas flotan en el espacio porque allí no existe *el aire* de la tierra
- Los astronautas pasaron muchos meses en el espacio, sin poder abandonar *el cohete* para nada
- Los niños hacen castillos de arena en *la playa* durante el verano.
- Los pescadores fueron recibidos por sus mujeres cuando el barco entró en *el puerto* de Santurce.
- Me quedo sin batería. Tengo el cargador conectado al móvil pero todavía me falta encontrar *una fuente* adaptable.
- Ni la fruta, ni la verdura, ni el pescado, lo único que le gustaba era *la carne* que le preparaba su madre.
- No había visto la señal y le pusieron una multa por cometer *una infracción* de tráfico
- No he podido leer el libro y mañana es el examen, puedes hacerme *una síntesis*, por favor?
- No paraba de llorar, pero el decía que sólo se le había metido algo en *el ojo* por el viento
- No podía entrar en el despacho porque había olvidado *la llave* para abrír
- No pudo seguir escribiendo en el dictado, porque se le había terminado *la tinta* del bolígrafo
- No sabe cocinar, entonces para celebrar San Valentín la invitó a cenar en *una habitación* de un hotel
- No todo es físico, una persona completa debe cultivar el cuerpo y *el espíritu* por igual
- Nunca mete un gol, no sabe jugar bien, pero le gusta mucho *la emoción* que le produce el fútbol
- Nunca sé dónde poner mi móvil y mi cartera, tengo que comprarme *una mochila* que combine con todo.
- Para cortar la carne se necesita *un cuchillo* de metal.
- Para ir a Barcelona no puedes coger ni el avión, ni el autobús, sólo tienes que coger *la*

dirección adecuada

- Para matar a un vampiro hay que clavarle una estaca en *la sién* con fuerza
- Para pedirle matrimonio se arrodillo ante ella y le dio *una joya* brillante
- Para ventilar la habitación mientras hace la cama, abre durante unos minutos *el ventanuco* estrecho
- Pepe era un chico muy activo, jugaba al futbol, al baloncesto, a la pelota, cada dia practicaba *un deporte* diferente
- Por haber robadovarias cosas en una tienda, pasará un año encerradoen *el penal* de Basauri
- Por precaución, siempre que cojas la moto debes ponerte *el casco* integral
- Prefiero que el té esté muy dulce, puedes pasarme *la miel* por favor?
- Sabíamos que estaba casada porque llevaba un anillo en *la cadena*, que nunca se quitaba.
- Se despertó sudando y temblando, había tenido *un sueño* terrible.
- Se estaba terminando el tiempo para patinar sobre hielo, iban a cerrar *el recinto* en diez minutos
- Si buscas a Ana, está preparando la comida, supongo que estará en *la cocina* preparándolo todo.
- Si no se afeita con regularidad le crece mucho *el pelo* y el bigote.
- Soy vegetariano, pero este jamón tiene muy buena pinta, voy a hacer *una excepción* esta vez.
- Su mujer se fue despues de 10 años de matrimonio. Eso le rompió *el corazón* en dos
- Subió la jaula del pájaro a un mueble temiendo que intentaría cazarlo *la mascota* de los vecinos
- Temía que si se portaba mal, al morir acabaría en *la miseria* seguramente
- Tenía mucha sed, pero por fortuna se topó en mitad de la calle con *un bar* abierto
- Tenía muchos cuadros, quería poner otro póster, pero no le quedaba ningún hueco en *el comedor* de su casa.
- Tenia solo un 5% de probabilidades de sobrevivir, pero él todavía era optimista, como dice el refrán lo último que se pierde es *la esperanza* de vivir
- Tiene un resfriado fuerte. El médico le dio una receta de antibióticos y un jarabepara *la tos* constante
- Todavía tenían un par de minutos de descuento antes de que el árbitro pitara *el final* del partido
- Todo quedó a oscuras porque se había ido *la luz* de pronto.
- Tuvieron que llamar a la grúa para que les remolcara *el coche* derrepente

- Uno de los ingredientes básicos de un bonbón es *la leche* según la receta
- Ya tenía puesto el traje, pero tuvo que llamar a su mujer para que le ayudara a atarse *el lazo* de la corbata

Experiment 2

LISTA 1	*LISTA 2*	*LISTA3*
CARNE	CLASE	COLCHÓN
CICATRIZ	COLIFLOR	ARAÑA
CRUZ	CRUZ	ESCALERA
FUENTE	FLOR	CAMELLO
HABITACIÓN	LECHE	GATO
HÉLICE	LLAVE	VIOLÍN
LECHE	LUZ	LOMBRIZ
NARIZ	MIEL	CLASE
NUEZ	MUERTE	LLAVE
PIEL	NARIZ	FLOR
PIRÁMIDE	NUBE	JAMÓN
RAíZ	PIEL	CUCHILLO
RED	PIRÁMIDE	MANIFESTACIÓN
SANGRE	RAíZ	CRUZ
SARTÉN	SANGRE	ELEFANTE
SERPIENTE	SERPIENTE	COLCHÓN
ÁRBOL	AGUACATE	RELOJ
AVIÓN	CACTUS	LECHE
CACTUS	CAFÉ	NARIZ
CAFÉ	COCHE	VACA
CAMIÓN	COHETE	MURCIÉLAGO
CHOCOLATE	CORAZÓN	MECHERO
COHETE	ELEFANTE	CAMALEÓN
CORAZÓN	JAMÓN	JAMÓN

DELFÍN	LIMÓN	PIRÁMIDE
GUANTE	MAÍZ	TIGRE
JAMÓN	MOTOR	QUESO
LÁPIZ	PEINE	MOTOR
PEINE	RELOJ	CALCETÍN
TIGRE	TIGRE	MAÍZ
TOMATE	TREN	OLLA
TREN	VIOLÍN	ARMARIO
BATERÍA	ARAÑA	ESPADA
CARTERA	BOTELLA	ANILLO
CORONA	CORONA	DELFÍN
ESCALERA	CUCHARA	CARTERA
GUITARRA	ESCALERA	NUBE
LAVADORA	GALLINA	GALLINA
MANZANA	LAVADORA	CARAMELO
NARANJA	MOCHILA	GAMBA
OLLA	MUÑECA	MOCHILA
PISTOLA	NARANJA	MANDO
PLUMA	OLLA	VASO
RAQUETA	PISTOLA	COHETE
RUEDA	PUERTA	RAQUETA
SILLA	SILLA	CERDO
TAZA	VACA	PIÑA
VENTANA	VENTANA	PERRO
ANILLO	ANILLO	PIEL
CABALLO	BOLÍGRAFO	GUITARRA
CEPILLO	BOLSO	CANDADO
CERDO	CEREBRO	PEINE
CEREBRO	CLAVO	TIGRE
CLAVO	CUCHILLO	PISTOLA
HUEVO	GATO	TREN
LIBRO	LIBRO	CIUDAD
MANDO	MANDO	CABALLO
OJO	MECHERO	SANDIA
PÁJARO	OSO	MIEL

PERRO	POLLO	TELÉFONO
PIANO	PULPO	MANIFESTACIÓN
POLLO	QUESO	CACTUS
PULPO	TELÉFONO	NAVE
QUESO	VASO	COLIFLOR
CICATRIZ	CICATRIZ	MELÓN
COLIFLOR	SARTÉN	CAMIÓN
SARTÉN	HABITACIÓN	LECHE
HABITACIÓN	SERPIENTE	CAMALEÓN
LUZ	NUEZ	HÉLICE
MUERTE	RAíZ	CALCETÍN
NUEZ	RED	CACTUS
RED	SANGRE	SILLA
CARNE	CARNE	AGUACATE
CLASE	CRUZ	TIBURÓN
FLOR	FUENTE	OLLA
FUENTE	HÉLICE	CLAVO
HÉLICE	LECHE	LAVADORA
LLAVE	NARIZ	COHETE
MIEL	PIEL	CAFÉ
NUBE	PIRÁMIDE	ELEFANTE
AGUACATE	ÁRBOL	SARTÉN
ÁRBOL	AVIÓN	CUCHARA
AVIÓN	CACTUS	SEMÁFORO
CAMIÓN	CAFÉ	LIMÓN
CHOCOLATE	CAMIÓN	BOLÍGRAFO
COCHE	CHOCOLATE	PISTOLA
DELFÍN	COHETE	SEMÁFORO
LIMÓN	CORAZÓN	CANDADO
ELEFANTE	DELFÍN	NARANJA
GUANTE	GUANTE	LAVADORA
LÁPIZ	JAMÓN	GUANTE
MAÍZ	LÁPIZ	NARIZ
MOTOR	PEINE	AVIÓN
RELOJ	TIGRE	CARNE

TOMATE	TOMATE	RAíZ
VIOLÍN	TREN	AGUACATE
BATERÍA	BATERÍA	CALLE
BOTELLA	CARTERA	PIÑA
CARTERA	CORONA	MUERTE
CUCHARA	GUITARRA	MECHERO
GUITARRA	LAVADORA	NARANJA
MUÑECA	NARANJA	OSO
PLUMA	PLUMA	HUEVO
PUERTA	SILLA	BOLÍGRAFO
TAZA	TAZA	BOLSO
ARAÑA	VENTANA	TIBURÓN
GALLINA	ESCALERA	LÁPIZ
MANZANA	MANZANA	HABITACIÓN
MOCHILA	OLLA	LIMÓN
RAQUETA	PISTOLA	PIRÁMIDE
RUEDA	RAQUETA	CHOCOLATE
VACA	RUEDA	PULPO
BOLÍGRAFO	CABALLO	CABALLO
BOLSO	CEPILLO	VIOLÍN
CABALLO	CERDO	MEJILLÓN
CEPILLO	CEREBRO	FUENTE
CERDO	CLAVO	COCHE
OJO	OJO	SERPIENTE
OSO	PERRO	MEJILLÓN
PERRO	POLLO	ÁRBOL
CUCHILLO	PULPO	ARMARIO
GATO	ANILLO	ESCALERA
HUEVO	HUEVO	VENTANA
MECHERO	LIBRO	CORAZÓN
PÁJARO	MANDO	GATO
PIANO	PÁJARO	SILLA
TELÉFONO	PIANO	COCHE
VASO	QUESO	GUITARRA
CRUZ	CARNE	ESPADA

FLOR	FLOR	PUERTA
LECHE	FUENTE	PATATA
LLAVE	HÉLICE	SARTÉN
MIEL	LLAVE	SERPIENTE
NARIZ	MIEL	LUZ
NUBE	NUBE	DELFÍN
PIEL	SARTÉN	PIEL
LUZ	LUZ	NUBE
PIRÁMIDE	HABITACIÓN	CORAZÓN
SERPIENTE	NUEZ	NAVE
COLIFLOR	COLIFLOR	CUCHILLO
RAíZ	CLASE	VENTANA
CLASE	MUERTE	HABITACIÓN
MUERTE	RED	LÁPIZ
SANGRE	CICATRIZ	CIUDAD
AGUACATE	AGUACATE	CRUZ
TREN	ÁRBOL	VASO
CACTUS	AVIÓN	BATERÍA
JAMÓN	TOMATE	CARAMELO
CAFÉ	LÁPIZ	MAÍZ
VIOLÍN	CAMIÓN	NUEZ
COCHE	CHOCOLATE	BOLSO
PEINE	VIOLÍN	ARDILLA
COHETE	GUANTE	RED
CORAZÓN	COCHE	SANGRE
MAÍZ	MAÍZ	CEREBRO
LIMÓN	DELFÍN	TAZA
MOTOR	LIMÓN	CICATRIZ
RELOJ	MOTOR	TREN
TIGRE	RELOJ	CLAVO
ELEFANTE	ELEFANTE	SANDIA
LAVADORA	GALLINA	CUCHARA
OLLA	RUEDA	CERDO
PISTOLA	RAQUETA	RAQUETA
ESCALERA	MOCHILA	BOTELLA

GALLINA	PUERTA	LLAVE
MOCHILA	MUÑECA	CHOCOLATE
VENTANA	BOTELLA	MOTOR
SILLA	VACA	RAíZ
PUERTA	MANZANA	TELÉFONO
MUÑECA	PLUMA	TAZA
BOTELLA	GUITARRA	POLLO
NARANJA	ARAÑA	GAMBA
VACA	CARTERA	BOTELLA
CORONA	CUCHARA	MANDO
ARAÑA	TAZA	CÁRCEL
CUCHARA	BATERÍA	ANILLO
PULPO	BOLSO	BATERÍA
BOLSO	VASO	CEPILLO
VASO	PIANO	RUEDA
MANDO	OJO	CORONA
QUESO	MECHERO	VACA
POLLO	PERRO	CICATRIZ
ANILLO	CABALLO	MELÓN
CEREBRO	BOLÍGRAFO	FUENTE
LIBRO	PÁJARO	MOCHILA
MECHERO	HUEVO	CEREBRO
BOLÍGRAFO	CERDO	GALLINA
CLAVO	CEPILLO	PUERTA
GATO	GATO	HUEVO
TELÉFONO	TELÉFONO	MUERTE
OSO	OSO	PEINE
CUCHILLO	CUCHILLO	LOMBRIZ
CALLE	CALLE	NUEZ
CÁRCEL	CÁRCEL	MIEL
NAVE	NAVE	RELOJ
MANIFESTACIÓN	MANIFESTACIÓN	RED
CIUDAD	CIUDAD	OSO
LOMBRIZ	LOMBRIZ	CARNE
MELÓN	MELÓN	AVIÓN

TIBURÓN	TIBURÓN	CÁRCEL
MEJILLÓN	MEJILLÓN	POLLO
COLCHÓN	COLCHÓN	ARDILLA
CALCETÍN	CALCETÍN	SANGRE
CAMALEÓN	CAMALEÓN	CARTERA
ARDILLA	ARDILLA	QUESO
PIÑA	PIÑA	PIANO
SANDIA	SANDIA	CORONA
PATATA	PATATA	PÁJARO
GAMBA	GAMBA	PLUMA
ESPADA	ESPADA	MURCIÉLAGO
SEMÁFORO	SEMÁFORO	PLUMA
CANDADO	CANDADO	PÁJARO
CAMELLO	CAMELLO	LUZ
CARAMELO	CARAMELO	GUANTE
ARMARIO	ARMARIO	HÉLICE
MURCIÉLAGO	MURCIÉLAGO	ARAÑA
CALLE	CALLE	PULPO
CÁRCEL	CÁRCEL	PATATA
MANIFESTACIÓN	MANIFESTACIÓN	MANZANA
CIUDAD	CIUDAD	OJO
LOMBRIZ	LOMBRIZ	CAMIÓN
NAVE	NAVE	MUÑECA
TIBURÓN	TIBURÓN	LIBRO
CALCETÍN	CALCETÍN	TOMATE
CAMALEÓN	CAMALEÓN	RUEDA
COLCHÓN	COLCHÓN	CEPILLO
MELÓN	MELÓN	COLIFLOR
MEJILLÓN	MEJILLÓN	PIANO
PATATA	PATATA	TOMATE
ARDILLA	ARDILLA	OJO
ESPADA	ESPADA	MANZANA
SANDIA	SANDIA	CAMELLO
PIÑA	PIÑA	ÁRBOL
GAMBA	GAMBA	CALLE

SEMÁFORO	SEMÁFORO	FLOR
ARMARIO	ARMARIO	CAFÉ
CARAMELO	CARAMELO	PERRO
CAMELLO	CAMELLO	MUÑECA
MURCIÉLAGO	MURCIÉLAGO	CLASE
CANDADO	CANDADO	LIBRO

Images used











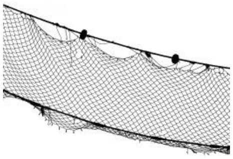
















































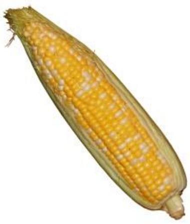






BIC



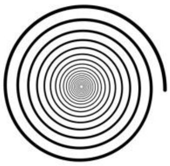















































YAMAHA















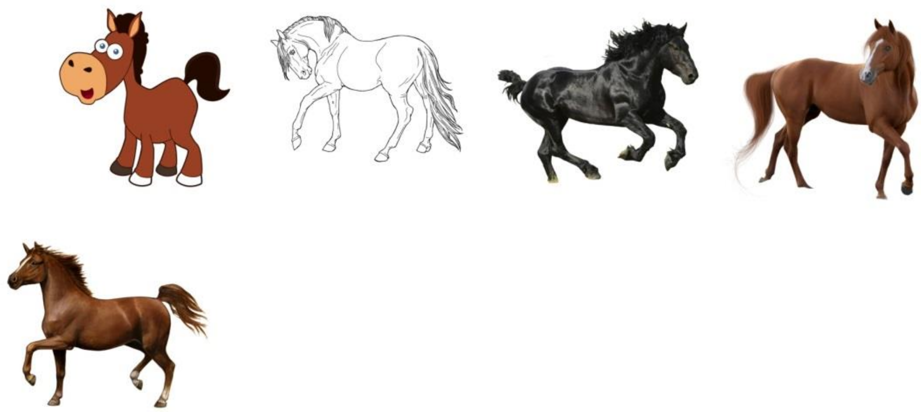


Experiment 3

Acordeòn

Aviòn

Caballo



Camara




Despertador





Flauta

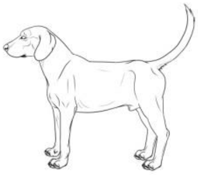




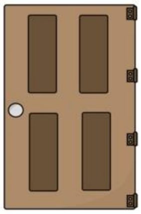
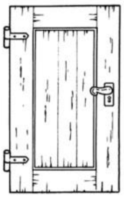
Pajaro

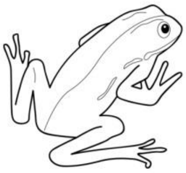
Perro

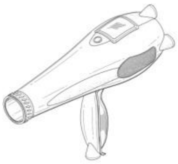
Puerta

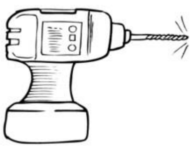
Rana


Secador


Taladro


Tijeras





Vaca