

Bangor University

DOCTOR OF PHILOSOPHY

Linguistic relativity electrified :|event-related potentials investigation of the way in which language affects cognition

Boutonnet, Bastien

Award date:
2013

Awarding institution:
Bangor University

[Link to publication](#)

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal ?

Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.



PRIFYSGOL
BANGOR
UNIVERSITY

Linguistic relativity electrified:

Event-related potentials investigation of the way in which language affects cognition.

Doctoral dissertation presented by:
Bastien Boutonnet

Supervisor:
Prof. Guillaume Thierry

Second Supervisor:
Dr. Panos Athanasopoulos

Chair:
Dr. Marie-Josèphe Tainturier

Submitted in partial fulfilment for the degree of Doctor of Philosophy.

Work completed in the School of Psychology, Bangor University

Acknowledgements

•

I would like to thank, first and foremost my supervisors Guillaume Thierry and Panos Athanasopoulos. Without Panos, I would never have discovered such a fascinating field of research nor would I have had the opportunity to come here to Bangor to learn and even better to contribute to the field's science. My work with Panos in the first year of my PhD was key in shaping my present research and for this I cannot be more thankful. Without Guillaume, I would never have been able to develop my burning interests in neuroscience when the time came to make a choice of whether to stay in Bangor or to move with Panos to Newcastle. I do not have words to express how grateful I am for Guillaume to have given me a chance when he barely knew me. The work in this thesis are the fruit of a perfect mix of circumstances along with an excellent supervision. Guillaume's never-fading excitement for research, his 100% overseeing of my work, his application of a distortion field capable of making me believe that I could achieve things I would have otherwise thought impossible have lead to an extremely educational, productive and thoroughly enjoyable supervisory experience.

•

The research in this thesis and the lightening-fast reconversion to neurophysiological techniques would not have been possible without the invaluable help from past and present colleagues in the lab. I would like to thank in particular Noriko Hoshino for having been so generous with her time in order to get me in the loop with stimuli presentation programming as well as hands on experience with data acquisition. To Jan-Rouke Kuipers and Ben Dering, I would like to show my appreciation, for their patience and useful technical expertise with aspects of the lab's management, and emergency technical help. To Clara Martin, the French super-woman, I would like to say thanks for taking some time out of her hectic two weeks of data collection in order to show me all stages of data processing for the first time and to just generally teach me how to be a superhero at collecting and analysing data at such a speed. This teaching has proven very useful when I visited her in Barcelona to collect data myself

•

I have had the immense pleasure to collaborate on a number of projects with amazing people. Some of these collaborations have allowed me to diversify my field of expertise and some of these have contributed directly to the work present in this thesis. In no particular order I would like to thank, Maria del Carmen Parafita-Couto, Clara Martin, Ben Dering, Manon Jones, Nestor Viñas-Guasch for being such great colleagues to work with.

•

I have been extremely lucky to receive funding for my PhD from the Bangor University School of Linguistics in my first year and then from the Economic & Social Research Council-sponsored Bilingualism Centre for the following years.

•

Outside of the lab, Bangor has proven an amazing place to spend 4 years. To the cost of some rather drastic burnouts I have had the immense pleasure to be part of so many student society with which I have been on stage acting, singing and dancing. I can safely say that those activities have been a great way to keep some grounding and always have a clear head about my own academic work. To Andrew Boorman, Joseph Hrycak, Tim Patten, Kayleigh Mobley, Katie Sims, Rachel Rivière, Rob Evans thank you for having been such special people. It has been great to be around you for however long or short, and you have successfully prevented me from becoming such a total social recluse.

•

Among those friends there are four people who have been too amazing not to deserve a few special words. Louise Ainsthworth, for being such an amazing human being. It is safe to say that without your friendship, support, giggles, and total ying-yang complementarity I would not have made it through these last years. Thank you so much and you better visit wherever I end up! Anna Beres, for being more than the new PhD student in the lab with whom I would have to share precious supervisory time and to whom I should be civil. Anna, you have become a great friend, and I am so thankful for your putting up with my write-up stress, for providing support when the mood was low and for sharing my fat intake from those Tuesday Texas BBQ Dominos pizza. I hope that I get to work with you again some time soon. To Kristin Koller, for having become such a great Brigantia friend and for sharing my fat intake on the Tuesday NewYorker pizzas. Gareth O'Neill, for having been one of the most important people in my life and for teaching me to believe in myself.

•

Finally, eternal gratitude and love goes to my family. Without them of this would not have been possible. I will be forever in debt for all the love, care they have given me as well as for all the sacrifices they have put up with so that I could get such an amazing education. Dominique, Alain et alexis, du fond de mon cœur merci pour tout

•

Linguistic relativity electrified:

Event-related potentials investigation of the way in which language affects cognition.

Contents

Abstract	9
Chapter 1	
General introduction and thesis overview	11
Chapter 2	
Categorisation, language & thought	19
<i>The brain as a categorisation device</i>	20
Cognition is categorisation	20
The brain's structure supports categorisation	22
<i>Language: a powerful tool for human categorisation</i>	26
Developmental and adult category learning	27
<i>Linguistic Relativity</i>	31
Colour perception and categorisation	34
Grammatical distinctions and cognition	40
Lexical items, object perception and categorisation	54
Strong effects that are weak, a problem?	57
<i>Interactive account of Whorfian effects: the label-feedback hypothesis</i>	60
Online and transient effects of labels on perceptual representations.	60
Computer model simulation of online label effects on perceptual representations.	61
Neural underpinnings of the hypothesis.	62
Implication for LRH research	63
Chapter 3	

Event-Related Potentials in language and perception 65

Event-Related Potentials, a particular analysis of EEG signal 66

EEG recording and setup 66

From EEG to ERPs 68

Waves with meaning: ERP components 70

The neural source of EEG signal 71

Advantages and limitations of ERP data 72

Overview of the use of ERPs in language processing and linguistic relativity 74

Meaning, relatedness and expectancy: N400 74

Grammar and syntax processing: P600 & LAN 75

Similarity, deviancy and language: Deviant-Related Negativities 77

Chapter 4

Unconscious effects of grammatical gender during object categorisation 79

Introduction 81

Methods & Materials 86

Results 90

Discussion 94

Conclusion 97

Chapter 5

Seeing objects through the language glass: Unconscious effects of labels on perception 99

Introduction 101

Materials & Methods 105

Results 109

Discussion & Conclusions 114

Chapter 6

Compound words prompt arbitrary semantic associations in conceptual memory 119

<i>Introduction</i>	121
<i>Methods & Participants</i>	124
<i>Results</i>	127
<i>Discussion</i>	129
Chapter 7	
General discussion	135
<i>Summary of the main results</i>	136
<i>Is the distinction between verbal and non-verbal processing meaningful?</i>	137
<i>Language is highly integrated in a not-so-hierarchical brain</i>	139
<i>On different levels of modulations</i>	141
<i>On late LAN interpretation</i>	142
<i>On Spreading Activation</i>	144
<i>On the absence of (Whorfian) behavioural effects</i>	145
<i>Online language effects on visual processing or effects of language on memory and decision processes?</i>	148
<i>Limitations and avenues for further study</i>	150
<i>Concluding Remarks</i>	153
References	155
Publications	187
<i>Boutonnet, B., Athanasopoulos, P., & Thierry, G. (2012)</i>	187
<i>Boutonnet, B, Dering, B., Nuñas-Guasch, N., & Thierry, G. (2013)</i>	187

Abstract

The possibility that language offers more to humans than the externalisation and transcription of thoughts has been at the heart of research on human cognition from millennia. In the past thirty years, the question of language-cognition interaction has received considerable attention in psycholinguistics and neuroscience, fuelling intense debate in several areas of human cognition, from basic brain organisation to social interaction. This thesis attempts to demonstrate the existence of effects of language on three levels of human cognition: perception, categorisation and conceptual associations. First, I show that distinctions brought about by grammatical gender influence object conceptualisation since gender information is retrieved spontaneously even when it is task-irrelevant. Event-related brain potentials collected in native speakers of Spanish asked to categorise triplets of pictures based on semantic category are significantly modulated by gender consistency between pictures. This effect is absent in native speakers of English, who do not have grammatical gender in their language. Second, I show that lexical distinctions in English that are not found in Spanish lead to differences in early visual discrimination. Native English speakers, who have two labels for the objects cups and mugs, perceive the two objects as more different compared to Spanish speakers who only have one word (*taza*). Third, language idiosyncratic lexical links brought about by word compounding result in predictable semantic associations between otherwise unrelated objects such as *neck* and *turtle* given their lexical link in compounds such as *turtleneck*. Overall, the work presented in this thesis offers novel evidence for linguistic relativity at an abstract grammatical level, at the interference between perception and conceptualisation and at the semantic/conceptual level.

Chapter 1

General introduction and thesis overview

One of the most distinctive features of human beings is their ability to communicate through language. Language has puzzled and excited intellectuals for millennia and continues to do so nowadays. Indeed, no communication system known to humans is as fascinating, complex, precise and powerful as language. By the “simple” utterance of sequences of sounds, language can refer to concrete and abstract entities in the world, externalise our thoughts, please or hurt people, make them laugh or cry, convince them to do things they would think impossible or discourage them from doing things that may be dangerous. Hearing /k/ followed by /ʌ/ and /p/ (cup) instantaneously evokes a series of representations such as a visual shape, the fact that it is used for drinking hot beverages like tea or coffee, and semantically associated representations such as a feeling of comfort after a day of work, gossiping amongst friends, and so on.

Language does not only label and refer to entities in the external world; it also allows people to talk about these entities, their relationships, and actions likely to affect them. Such advanced communication is made possible by a series of arbitrary sequences of sounds, which make up words arranged into sentences to form and communicate propositions, which are of essentially infinite diversity. In a nutshell, language sounds are to language what atoms are to matter; they enable humans to communicate unlimited information from a limited number of basic elements.

It is beyond discussion that language is an extraordinarily elaborate tool for the externalisation and internalisation of thought. However, is language a medium dedicated entirely and exclusively to the externalisation of thoughts?

Or, does language extend beyond the activity of speaking and reasoning?

Famously, for instance, Benjamin Lee Whorf –a student of Edward Sapir who worked as a fire insurance inspector– was drawn to the study of language when he was taken aback by the fact that “empty” fuel drums were perceived as less dangerous than full ones. In fact, these fuel drums were just as dangerous since they were “full” of highly flammable vapours. Did such a misconception merely derive from a lack of knowledge or did language use somehow change the conception of presence?

William James (1890) made an interesting observation regarding the effect that words have on perceptual experience:

“I went out for instance the other day and found that the snow just fallen had a very odd look, different from the common appearance of snow. I presently called it a ‘micaceous’ look; and it seemed to me as if, the moment I did so, the difference grew more distinct and fixed than it was before. The other connotations of the word ‘micaceous’ dragged the snow farther away from ordinary snow and seemed to aggravate the peculiar look in question.” (p. 512)

Here, and although it is derived from subjective perceptual experience, the act of labelling a particular property of something, seemed to change not only the concept of the object but even make its perceptual characteristics more salient.

Such questions have been at the heart of research on language and human cognition, a considerable amount of attention and research has been devoted to the understanding of whether and how language changes humans’ perception of the external world. In the past thirty years, reconceptualisations of the famous writings of Whorf (1956) have been put forward, calling for controlled

experimental procedures and recommending safe-guards for the interpretation of effects of language on thought (Gentner & Goldin-Meadow, 2003; Gumperz & Levinson, 1996; Hunt & Agnoli, 1991; Lucy, 1992a). On the experimental front, effects of language on other cognitive processes have been inferred from the observation of an influence of cross-linguistic differences on the expression of grammatical number, showing that languages that do not mark number on nouns but do so via classifiers lead their users to categorise objects on the basis of substance rather than shape (e.g., Athanasopoulos, 2006; Cook, Bassetti, Kasai, Sasaki, Takahashi, 2006; Lucy, 1992b; Saalbach & Imai, 2007; Zhang & Schmidt, 1998). In a similar vein, grammatical gender expression on noun has been shown to affect speakers' object representations in covert gender assignment tasks (e.g., Bassetti, 2007; Forbes, Poulin-Dubois, Rivero, & Sera, 2008; Kurinski & Sera, 2010; Sera, Berge, & Castillo-Pintado, 1994; Sera *et al.*, 2002), judgements and adjective-association tasks (e.g., Boroditsky, Schmidt, & Phillips, 2003; Phillips & Boroditsky, 2003), and priming paradigms (Boutonnet, Athanasopoulos, & Thierry, 2012; Cubelli, Paolieri, Lotto, & Job, 2011). Lexicalisation constraints on spatial representation and event conceptualisation (e.g., focus on manner vs. end-point of motion) seem to affect speakers' event description and recollection (e.g., Athanasopoulos & Bylund, 2012; Bowermann & Choi, 1991; Majid, Bowerman, Kita, Haun, & Levinson, 2004; Papafragou & Selimis, 2010) or to elicit different gaze patterns when exploring scenes depicting events (Flecken, 2010). Finally, differences in terminology have been shown to affect speakers' perception of colour in behavioural (Athanasopoulos, 2009; Franklin, Drivonikou, Clifford, Kay, Regier,

& Davies, 2008; Ozgen, 2004; Roberson, Davidoff, Davies, & Shapiro, 2005) and neurophysiological (Athanasopoulos, Dering, Wiggett, Kuipers, & Thierry, 2010b; Clifford, Holmes, Davies, & Franklin, 2010; Liu *et al.*, 2010b; Thierry, Athanasopoulos, Wiggett, Dering, & Kuipers, 2009) investigations.

Despite the growing body of evidence in favour of Whorfian effects, the dispute between researchers who try to provide experimental support and those who formulate the strongest and most provocative critiques (e.g., Pinker's (1995) dismissal: "no matter how influential language may be, it would seem preposterous to a psychologist that it could reach down into the retina and rewire the ganglion cells." p.63.) has not ceased. But, as the French saying goes, "*la critique est aisée mais l'art est difficile*" [*Criticism is easy but the art is difficult*]. When they show a minimum amount of consideration for the evidence accumulated to date, critics of linguistic relativity ridicule the data as fuzzy support for a "weak" version of the hypothesis. It is undeniable that a substantial part of the experimental literature available to date fails to provide convincing evidence for strong Whorfian effects and that many papers reached mixed conclusions. But then, this should be considered a wealthy situation, for positive scientific evidence accumulating always in the same direction, is a sign of experimental bias and selective reporting. The camp of the critics has an easy game, but "*l'art est difficile*". Because we should be clear, the task is very hard. Establishing spontaneous effects of language on other cognitive processes in well-controlled experiments, especially without trivially involving language processing, that is, directly targeting "strong" Whorfian effects rather than "trivial effects of language on cognition" (Pinker, 1995) is a major undertaking.

This, however, is what I have attempted to do in the experimental work reported in this thesis. In line with previous attempts to demonstrate Whorfian effects at the perceptual level, the thesis aims to extend previous findings in the domains of grammatical gender, lexical distinctions, and language-specific terminology, in line with Lupyan's (2012) recent theoretical and experimental developments. The three experimental chapters investigate Whorfian effects at three of the main levels of human cognition: perception, categorisation and conceptual associations.

First, I sought to show that distinctions brought about by grammatical gender, in a language like Spanish, lead to differential and spontaneous recollection of gender information, even when such information is entirely irrelevant. Triplets of pictures either related semantically or not were presented to speakers of Spanish and English who were asked to perform a semantic categorisation task. Unbeknownst to them, the gender of the picture names was manipulated orthogonally so that, independently of semantic relatedness, the gender of the last picture either matched the gender of the two previous ones or not. I identified an effect of gender congruency in a task where such a feature is irrelevant and unnecessary. The results suggest that gender information is spontaneously retrieved and that the thought-content is modified online, such that gender information becomes integral to object representation.

Second, I show that lexical distinctions in English that are not found in Spanish, lead to differences in early visual discrimination. I measured differences in brain responses elicited by pictures of a cup and that of a mug in native speakers of English, who have a different label for the two objects, and in

native speakers of Spanish, who have only one label to refer to these same objects (*taza*). The use of an oddball paradigm similar to that used by Thierry *et al.* (2009) enabled us to study an early modulation of brain activity thought to be mostly automatic and highly diagnostic of basic visual perception. As predicted by the label-feedback hypothesis (Lupyan, 2012), early differences were found in the native speakers of English, but not in the native speakers of Spanish, suggesting that the label distinction in English lead to differences at the interface between perception and conceptualisation.

Third, I set out to determine whether relatively artificial links at a lexical level, brought about by word compounding, between objects that are otherwise unrelated (e.g., *turtle* and *neck* in *turtleneck*) would result in predictable semantic associations. This experiment investigated linguistic relativity at a level more conceptual than perceptual, since the pictures used were *completely* unrelated perceptually. I chose to present pictures in pairs which were either related semantically or not and of which the labels could form part of a compound word or not. Brain activity seemed to show no effect of compounding in terms of semantic relatedness, except when the pictures were presented in the reverse order of that of compounds (e.g., picture of a human neck – picture of a turtle). This result suggests that concepts, arguably unrelated, become related at a conceptual level because of the existence of a compound word, which has fostered an artificial link between lexical entries.

Overall, I believe my work presents novel evidence for linguistic relativity at an abstract grammatical level, at the interface between perception and conceptualisation and at the semantic/conceptual level.

Chapter 2

Categorisation, language & thought

1. The brain as a categorisation device

1.1. Cognition is categorisation

In the previous section we emphasised the importance of language as a faculty definitional of humans. Here, however, we will take a few steps back and consider an essential feature of human cognition: categorisation. Let us consider the following example, (cf. Harnad, 2005). First, Borges' (1962) fictional character Funes, who, as a consequence from falling off his horse, acquires a disorder that allows him to have unlimited rote memory. What this meant was that he would no longer forget anything, every instant of his life would be stored forever and he could, very much like one would from a DVD, "replay" those stored experiences. Funes was therefore puzzled by the fact that people around him referred to a particular dog, present at a particular moment, in a particular place and position, by the same name as they called it at another moment, in a different place, time and position. His infinite memory, which we could intuitively think of as an advantage, was in fact a serious handicap since, for him, every instant was stored as unique, which made all situations ultimately incomparable. Indeed, his inability to forget would mean that he would not be able to abstract over a number of occurrences and remember the essential invariants of entities (i.e., generalise and categorise). In fact, he should not have been able to grasp the concept of a dog given the fact that every single instant would be stored as an individual and particular "entry". Funes would not be able to recognise a particular dog as the same dog in a different situation or even the fact that it is a dog at all. In fact, Funes would not be able to do much apart from being a living collection of instants witnessed from his own perspective.

Borges himself pointed out the need to categorise in these terms:

“Pensar es olvidar diferencias, es generalizar, abstraer.

En el abarrotado mundo de Funes no había sino detalles, casi inmediatos.”

Thinking is forgetting differences, generalising and abstracting.

Funes’ crowded world was almost solely made of immediate details. (My translation)

Borges, Funes el memorioso.

Although we considered an extreme case, such a situation describes many cases of people with over developed rote memory (the case of Luria’s (1986) patient “S”) or in fact, many of the characteristics of disorders like autism. Those cases should put in perspective the importance of categorisation on cognition – a process one can often take for granted given its automaticity and apparent easiness. Categorisation is essential for cognition because it enables us to build our semantic memory and to interact with the outside world. By extracting essential characteristics of animals and artefacts from the outside world (e.g., physical resemblance, sounds, origin, matter, use etc.) and by grouping those into categories based on the links between one another, we can construct networks of knowledge of how to interact with objects independently of us directly interacting with them in the present moment. Categorisation provides the gateway between perceptual experience and cognition and it occurs in all sensory modalities. People categorise sounds, smells, tastes, physical appearance, tactile sensations and even their own emotions and thoughts. The representations made available from categorisation become central for future

cognitive processing, where they are combined with other representations and can trigger other processes such as the intention to achieve a goal and one could argue that most cognitive processes, in fact, begin with some form of categorisation (Barsalou, 1992). As Harnad (2005) puts it: “To Cognize is to Categorize”. Although, these claims may sound extreme, given the fact that we are not necessarily conscious that we are categorising, and that we seem to be able to do it so effortlessly, the central role of categorisation should not be downplayed. In fact, it is usually those processes that we take for granted because we hardly notice them occurring, that are essential. Perhaps one of the reasons they are so effortless is because they are so fundamental and well integrated into our organisms. Indeed, categorisation is also highly supported by the brain’s organisation, as we discuss in the following section.

1.2. The brain’s structure supports categorisation

The idea that categorisation is essential for cognition does not only hold on the philosophical standpoint, much like in the way we have talked about it so far, but the idea that the brain is indeed a machine well prepared to support categorisation, is also evidenced by neurophysiological studies, which have isolated areas and mechanisms that are central to the emergence and learning of categories. Specific areas as well as structural connections, and phenomena like plasticity have all been related to category learning. Part of the visual system, the inferior temporal cortex (ITC) features neurons that have complex shape selectivity, which may have category-like tuning (Desimone, Albright, Gross, &

Bruce, 1984; Logothetis & Sheinberg, 1996; Tanaka, 1996), suggesting that, as early as the visual system, the notion of category is already present.

In a less clear cut manner than category-selective neurons, areas of the brain have been identified in category learning with some areas sensitive to features, experience and variability; and some areas, associated with processes such as decision-making, rule and criterion-based learning. While the ITC is mainly devoted to the processing of perceptual features and similarities of stimuli, the prefrontal cortex (PFC) has been strongly associated with the representation of boundaries or conjunctions between stimuli (Jiang, Bradley, Rini, Zeffiro, Van Meter, Reisenhuber, 2007; Li, Mayhew, & Kourtzi, 2009) as well as abstract rule-based categorical distinctions (Muhammad, Wallis, & Miller, 2006). The medial temporal lobe (MTL) has also reliably been identified in category learning, particularly instance-based learning, storing regularities and exceptions to rules (Love, Medin, & Gureckis, 2004).

Still from a structural point of view, corticostriatal loops (close anatomical loops involving the cortex and the basal ganglia, BG) seem to be of prime importance in the formation of categories. Corticostriatal loops (cf., **fig. 1**) form channels within the BG that return inputs to the same cortical areas that gave rise to their initial input (Hoover & Strick, 1993; Kelly & Strick, 2004) and enable recursivity where the results from one iteration are fed back through the loop for further elaboration. It has been shown that all four loops can be recruited during categorisation learning (Seger, 2008), which is consistent with the type of processes required during categorisation: processing of visual

stimulus, preparing and executing a motor response, receiving and processing feedback (Seger & Miller, 2010).

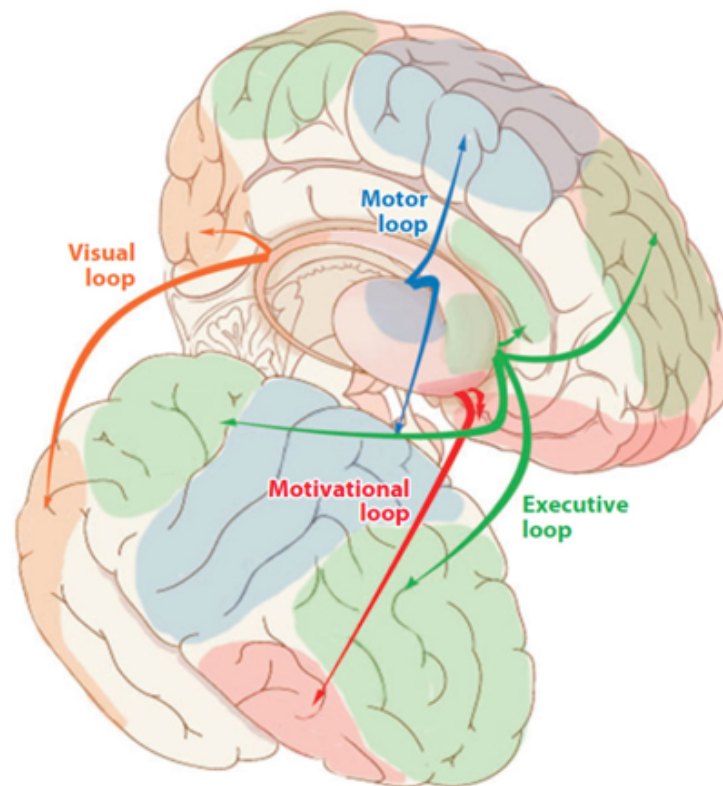


Figure 1. Illustration of corticostriatal loops from Seger & Miller, 2010 (p.209). The motor loop connects the motor cortex with the posterior putamen. Executive loop connects the prefrontal cortex and the parietal cortex with the anterior caudate nucleus. The motivational loop connects the ventral striatum with the orbitofrontal cortex. The visual loop connects extrastriate and inferotemporal cortices with the posterior caudate nucleus. (Reproduction of material granted by the authors).

Finally, different cortical regions have different plasticity rates (fast subcortical plasticity vs. slower cortical plasticity). This phenomenon is also thought to be of crucial importance in categorisation. Fast plasticity allows large changes in synaptic weights with each episode, which lead to rapid storage and quick learning. Slow plasticity, on the other hand, induce smaller changes, which are not tied to specific events and require multiple exposures to emerge. This means that regions with slow plasticity will be influenced by a sort of

common average of stimuli: generalisation (McClelland, McNaughton, & O'Reilly, 1995). Seger & Miller (2010) suggest that recursive interactions between the BG's fast plasticity and slow cortical plasticity underlie the construction of categories and allow abstractions. Imbalance between fast and slow plasticity could cause the BG to be stronger than normal and details might overwhelm the cortex, much like it is the case in autism or in the fictional case of Funes (Borges, 1962). In fact, recent work has shown that many psychiatric and neurological symptoms such as autism, could be explained by an imbalance between slow and fast plasticity, where, due to a disruption of glutamate receptor mGluR5, dopamine is in abnormal abundance, which results in faster-than-normal plasticity in striatum making generalisation almost impossible (Dölen *et al.*, 2007).

This brief overview of the neurophysiological substrate for categorisation (see Seger & Miller, 2012 for a more exhaustive review) should have made it clear that the brain is strongly prepared for the acquisition and development of categories in a very dynamic, distributed and therefore flexible manner. It should be noted, however, that these processes are not unique to humans but nonetheless necessary for the advanced level of categorisation implemented.

At this point, one might ask where is language in all this? How can human cognition be so much more advanced than the one of other animals with advanced cognitive abilities like other mammals and primates in particular? Is this advantage just due to better and more efficient brain organisation or is there another set of skills augmenting it? Language, a definitely human skill, is probably where a significant part of this advantage resides. In the following

section, we argue that human language and categorisation are intimately linked, which would at least in part explain the power of categorisation in humans.

2. Language: a powerful tool for human categorisation

In this section, I will present arguments which support the idea that language is not just a tool for the externalisation and exchange of thought but rather that it considerably augments our categorisation capacity. Language is characterised by a finite, albeit large, set of items. Despite this apparent limitation, an infinite number of utterances can be produced. In fact, this finite-to-infinite relationship in itself justifies the need for generalisation, abstraction, and therefore categorisation. To utter or hear the word “toaster” is in itself categorisation, for this label invokes all the possible specific instances of toasters already encountered or even to be encountered and with it a series of representations such as its use in toasting bread, the fact that it typically has two slots, a thermostat, belongs in the kitchen and so on.

We can, in fact, say that language provides an external manifestation of internal categorisation processes. Indeed, linguistic labels, by definition, have to capture the necessary differences and leave out irrelevant ones. In other words, language provides “pre-cooked”, ready-to-(eat)use categories, which language-able organisms can latch on to and which helps bootstrap natural diversity into palatable semantic bites. It is, therefore, not surprising that language is considered to have significant advantages on the development of human cognition and even evolution (Harnad, 2005). Indeed, Cangelosi & Harnad

(2001) present simulation data, which highlight considerable advantages of learning categories through “linguistic theft”, as they call it, rather than through sensorimotor implicit toil.

2.1. Developmental and adult category learning

Moving away from the evolutionary perspective, linguistic theft (Harnad, 2005) may not require simulations to be put to the test. In fact, such phenomena have been widely reported by developmental studies that have investigated the consequences of language on categorisation and category learning. Firstly, language, and particularly labels, grabs infants’ attention when used to determine objects in early development. Baldwin and Markman (1989) show that, as early as 10 month-old, infants devote more attention to objects that have a label than to those which do not. This not only happens during naming but also post-naming where infants gaze longer at the target toys that were labelled in previous tasks. Similar evidence has been consistently replicated by Waxman and colleagues (Balaban & Waxman, 1997; Waxman & Booth, 2003) in 9- and 11-month-old infants where infants’ performance significantly improved when objects were labelled than when they were not. Xu (2002) reports an even deeper effect of language in a similar age group, with 9-month-old infants in an object individuation task, where infants’ ability to conceive of two objects as different was helped by language compared to conditions where infants had been familiarised to objects presented in conjunction with sounds or emotional stimuli. Additionally, when infants reach a certain threshold in word-learning (around 14 months) it has been shown that they build more specific

expectations depending on the word class of the label: nouns seem to guide infants towards a categorical level exclusively, whereas adjectives guide them towards both category- and property-based commonalities (Waxman, 2001; 1999). Several studies have also reported a beneficial effect of language on the acquisition of abstract categories such as spatial relations (Casasola, 2005; Pruden, Roseberry, Göksun, Hirsh-Pasek, & Golinkoff, 2012).

It is important to note a few concerns, however, with the interpretation of the previously reviewed data since they rely on the study of common categories such as cars, horses, fruits (and variations of their sub- or super-ordinate levels), with which infants were undeniably already familiar. It is unclear whether labels resulted in the teaching of new categories or whether they merely reactivated categories that were previously acquired (Plunkett, Hu, & Cohen, 2008). Furthermore, it has been demonstrated that children exhibit out-of-category novelty preferences in the absence of any guidance from labelling (Behl-Chadha, 1996; Eimas & Quinn, 2008; Younger, 1985). That is to say that in paradigms which measure looking time, children look longer at the items with which they are not familiar as opposed to the ones with which they have received training (Behl-Chadha, 1996). Therefore this research suffers from a lack of baseline in which no labelling is taking place, since previous studies have not always controlled the amount of exposure to experimental stimuli in relation to the use of labels (see Plunkett *et al.*, 2008, for an extensive critical evaluation of the literature).

These limitations can however be resolved by adopting a “process-oriented” approach such as that proposed by Oakes and Madole (2000), where a no-sound

baseline is introduced and in which completely novel stimuli are used that do not correspond to potentially pre-existing categories in the infants' conceptual system. This is exactly what Plunkett *et al.* (2008) did in an experiment where they familiarised infants with pictures of cartoon-like characters which varied on dimensions such as tail thickness, leg length and ear orientation. The familiarisation phase was meant to induce category learning such that drawings would be grouped in two categories: one consisting of low values of tail thickness, leg length and ear orientation and the other consisting of combinations of high values of such parameters. This phase replicated Younger's (1985) results showing that infants had learned the categorical distinction they were familiarised with. More to the point, the authors carried out a series of experiments manipulating label-category mappings. When the two labels were correlated with visual categories, infants learnt those just as well as they did in the no-label condition. However, uncorrelated labels disrupted the formation of any category and the use of one single label across both categories lead to the formation of one all-encompassing category. These results therefore demonstrate a link between labels and categories where labels can strongly influence category formation in infants as young as 10-month-old.

The data reviewed in this section provide significant evidence suggesting that language can and does play a crucial role in human cognition in the earliest stages of language acquisition. Cognition is dramatically dependent on categorisation, which is itself strongly tied to language.

To test the effective power of language on categorisation in adults, Lupyan, Rakinson, and McClelland (2007) designed a task, which was meant to

investigate whether being trained on category learning via linguistic labels compared to facts had a better impact on categorisation. The task required participants to learn to classify 16 “aliens” into a group that is safe to approach and another to be avoided. The stimuli involved subtle perceptual differences in the configuration of the head and body of the creatures. The training phase involved conditions where participants received a label together with the stimuli and another condition where no label was provided. The results show that subjects who received labels during the training generalised what they had learned to new stimuli presented in the test phase and performed better overall than subjects who had not received label information. An additional experiment involved location information about the aliens such as “lives above” or “lives below” instead of linguistic labels. In this version of the experiment, receiving cues about location (although matched in the same manner as the linguistic labels in the first experiment) did not lead to a performance benefit. These results provide strong evidence that language is a powerful tool for categorisation.

Coming back to the series of questions we asked at the end of the previous section, we have now gathered a pretty good insight into the power of language for categorisation. From this insight, one of the further questions we could ask is how the language advantage for categorisation is established in the human brain, that is, whether language is peripheral to categorisation or part and parcel of the categorisation process.

3. Linguistic Relativity

The view that language may be transformative of cognition and perception, and offer more than “simple” communication, which has commonly been referred to as the Sapir-Whorf hypothesis and related to the writings of Benjamin Lee Whorf (1956) was not incepted in the 1950’s. It has been the matter of long-standing interest from many philosophers like Plato, Kant and Müller and it is still the matter of debate nowadays in anthropology, linguistics, psycholinguistics and neuroscience. Insights into this question are of central importance and have far-reaching implications for any understanding of human cognition, from functional brain arrangement (modularity and encapsulation [Fodor, 1975; 2008] vs. distributed networks and interaction [Elman, 2004; Kiefer & Pulvermüller, 2012; Mahon & Caramazza, 2008; 2009; McClelland & Rumelhart, 1981; Pulvermüller, 1999; 2012]) to social interaction. Early ideas on the interaction between language and thought have often conflated the two, where thinking is equated to language or as Müller (1909) put it: “language is identical with thought” (p. ii). Such conflation leads to the interpretation that thought is not possible without language: the most deterministic view of the language-thought relationship. Whorf’s writings, which are most famously associated with this question, have often been considered to argue in favour of such a deterministic position. Indeed, Whorf’s famous quote can easily be taken to mean that language deeply *defines* our perception and categorisation of the world:

“We dissect nature along lines laid down by our native languages. The categories and types that we isolate from the world of phenomena we do not find there because they stare every observer in the face; on the contrary, the world is presented in a kaleidoscopic flux of impressions which has to be organized by our minds –and this means largely by the linguistic systems in our minds ... Language is not simply a reporting device for our experience, but a *defining* framework of it (Whorf *et al.*, 1956, p. 213).” (My emphasis)

Other times, Whorf writes about the language-thought interaction in a very similar way to modern connectionist accounts of cognition, which advocate *interactivity* between language and thought, rather than *definition*:

“Any activations [of the] processes and linkages [which constitute] the structure of a particular language ... once incorporated into the brain [are] all linguistic patterning operations, and all entitled to be called thinking.(Whorf, 1937, p. 57–58, cited in Lee, 1996)”

Clearly, and even at the time Whorf wrote about this hypothesis, we can see that he had already departed from linguistic determinism although it is precisely what critics and proponents of modular theories of mind (Chomsky, 1965; 2000; Fodor, 1975; Pinker, 1995) condemned. This view is well expressed by Pinker (1995, p. 65): “Most of the experiments have tested banal ‘weak’ versions of the Whorfian hypothesis, namely that words can have some effect on memory or categorization ... but that is hardly surprising.” This view restricts the interest of the Whorfian hypothesis to the measurement of “strong” effects

and therefore its deterministic interpretation. However, in light of compelling evidence that infants and non-human primates are capable of entertaining relatively high-level thinking without language (Feigenson, Dehaene, & Spelke, 2004; Gallistel, 1989; Hare, Call, & Tomasello, 2001; Penn, Holyoak, & Povinelli, 2008; Phillips & Santos, 2007) such a deterministic standpoint becomes untenable. Does this mean, then, that language should not affect thought at all? The past twenty years of research clearly disprove this view, since they have been characterised by a flurry of reinterpretations of the hypothesis, tinkering between “strong” and “weak” readings of what has become known as the Linguistic Relativity Hypothesis (LRH). This tinkering gives an impression of a field with split personalities, which may have some consequences on the ability of researchers to make clear predictions and to scientifically put them to the test. This problem, we will argue, is mainly due to methodological limitations and failure to integrate the Whorfian effects within a psychologically sustainable theory (see **section 4**).

During the 1990s revival of the LRH, scholars such as Lucy (1992a; 1992b), Levinson (Gumperz & Levinson, 1996) and Slobin (1996) have provided theoretical and methodological frameworks of testing. They put a strong emphasis on using experimental methods and designs that tap both into linguistic and non-linguistic processes so as to get an idea of the magnitude and nature of the influence of language on other cognitive processes. However, these authors did not necessarily share expectations as to what these effects might be and the field has been crippled by constant inability to reconcile “weak” and “strong” effects.

A perfect example of this problem is reflected by Slobin's (1996) strongly limitative conception of the hypothesis, which contends that influences of language on thought are restricted to thought formulated *during* and *for* language use: *thinking for speaking*. In that regard, speakers of a particular language may, at the time of language production, be more attuned to distinctions made in that language. Therefore, speakers of French, a language that specifies tenses on verbs (past, present and future), would pay more attention to the moment events occur as opposed to speakers of Chinese who do not mark tense on verbs and that the influence would be limited to conditions where speech is produced or to be produced. Although Slobin's *thinking for speaking* hypothesis is highly compatible with most of the empirical data to date, this is actually its limitation since it does not have a diagnostic value to assess weak versus strong interpretations of the LRH.

3.1. Colour perception and categorisation

All humans with normal colour vision share the same physiological basis of colour perception resting on the presence of three colour receptors –blue, green and red spectrums– (Jordan & Mollon, 1997). Yet there are differences in the way languages describe this continuum, although arguably the physical nature of the signal is the same whether one is a speaker of English or Greek. Some languages describe the whole spectrum with only two colour terms, and most of the cross-linguistic variation falls between three and eleven (Berlin & Kay, 1969) with some language like Turkish (Ozgen, 2004), Russian (Davies & Corbett, 1997), Greek (Athanasopoulos, 2009), or Japanese (Uchikawa & Boynton, 1987)

having additional categorical boundaries, for example, in the blue domain. Not only does the number of basic colour terms vary from one language to another, the nature of colour boundaries can also vary markedly.

Unsurprisingly, research in the domain of colour terminology, and its effect on perception and categorisation, has generated a great deal of interest (humans like to debate colours as illustrated by the French saying: “*Les goûts et les couleurs ne se discutent pas*” [There is no accounting for tastes and colour]). Indeed, it is in this domain that the most controversial and opposed claims about the influence of language on thought have been made: from the most nativist to the most deterministic. This is perhaps best reflected in the work of Berlin & Kay (1969) who originally proposed a theory of colour universals limited to 11 colour terms and contended that language followed constrained paths in their evolution of colour terms. Recently, however, the same authors have come to admit that their original claims were, at least partly, wrong. One of the first challenges to the theory stems from Özgen and Davies’ (1998) study of Turkish colour terms which identified the possibility that Turkish had an additional basic colour term in the blue spectrum. A series of behavioural cross-linguistic studies (Davidoff, Davies, & Roberson, 1999; Davies & Corbett, 1997; Roberson, Davidoff, Davies, & Shapiro, 2005; Winawer, Witthoft, Frank, Wu, Wade & Boroditsky, 2007) have reported effects of colour terminology on memorisation and categorisation of colours, which suggest that participants are more likely to group or judge two colours to be similar if they share the same label, disregarding objective perceptual distance. Crucially, in a longitudinal study, comparing English to Himba in 3- to 5-year-old children, Roberson *et al.* (2004)

showed that, as they start developing their colour vocabularies, children's performance on a series of memory tasks start diverging towards the distinctions present in their particular languages, whereas children who had not yet acquired colour terms make similar memory errors in both language groups, i.e., based on perceptual distance rather than any particular set of "innate" categories.

Evidence that Whorfian effects may happen in early stages of processing has been obtained in more controlled experimental designs. Winawer *et al.* (2007) compared English and Russian speakers' performance on a speeded similarity judgement task, where participants were presented with triads of blue squares varying in degrees of lightness and had to decide whether the target square matched the left or right square. Pairings were arranged so that the mismatching colour was either close or far to the target, in perceptual space. Most importantly, for the Russian speakers this meant that sometimes all three colours would fall under the same label, whereas in other conditions the mismatching colour was described by another label. Their results show that Russian speakers' reaction times (RTs) were significantly faster in the between-category trials although this was not the case for English speakers for whom all 20 shades of blue still fall within the same linguistic category.

A series of studies have also found strong evidence of early influence of linguistic labels in more perception-centred and controlled within-subjects paradigms (Drivonikou *et al.*, 2007; Gilbert, Regier, Kay, & Ivry, 2006; Lu, Hodges, Zhang, & Wang, 2012). For example, Gilbert *et al.* (2006) used a visual search paradigm, a type of task in which the measured RTs are strongly

diagnostic of visual processing. Participants were presented with an array of colour chips arranged in a circle around a fixation cross. The coloured chips were either two members of the blue or of the green category. The arrangement was so that the standard and deviant chips either all belonged to the green or blue category (within-category condition) or the deviant belonged to the opposite category (between-category condition). The authors took care of maintaining a similar perceptual distance between the within- and between-category conditions so that the only modulation would come from the linguistic label (for a detailed discussion see note “§§”, p. 493 in Gilbert *et al.*, 2006). On each trial participants were asked to indicate whether the deviant coloured chip was in the left or right half of the circle, by making speeded keyboard responses. Participants were significantly faster in between-category conditions and this significant main effect interacted with visual field so that between-category judgements were faster when the deviant was presented in the right-visual field (i.e., processed by the left hemisphere). The results therefore show a significant influence of linguistic labels on colour perception with the added neural dimension that such an effect may be supported essentially by the left hemisphere, known to have a key role in language processing. Similar conclusions were also reported by visual search experiments carried out on children populations (Daoutis, Franklin, Riddett, Clifford, & Davies, 2010; Franklin *et al.*, 2008). Although these data lend strong support in favour of Whorfian effects they do not allow the authors to conclude whether these effects stem from a “direct” impact of language on visual processing or whether the RT

benefits are a consequence of postperceptual (strategic) processes where language influences the decision at a late response stage.

Neurophysiological investigations on colour perception provide the clearest and most compelling data available to date supporting an effect of labels on colour perception. Using similar visual search or oddball paradigms while recording event-related brain potentials (ERPs), several studies have reported early and robust effects of linguistic labels on basic visual processing (Athanasopoulos *et al.*, 2010b; Clifford *et al.*, 2010; Clifford, Franklin, Davies, & Holmes, 2009; Liu *et al.*, 2010b; Thierry *et al.*, 2009). Thierry and colleagues (2009) were the first to report early effects of colour terminology on visual processing. They recorded ERP correlates of colour change detection in Greek-English bilinguals who have two colour terms for blue (*ble* – ‘dark blue’ and *ghalazio* – ‘light blue’) and a control group of English monolinguals. They found that native speakers of Greek exhibited a greater visual mismatch negativity (vMMN) elicited by blue deviants than English controls. These results therefore confirm and extend the Russian data from Winawer *et al.* (2007) discussed earlier. Several subsequent studies have also refined these findings by looking at lateralisation (Clifford *et al.*, 2009; 2010; Q. Liu *et al.*, 2010b) and have confirmed behavioural data looking at visual fields with neurophysiological correlates. Data from bilingual populations have led to an even deeper insight into the plasticity of Whorfian effects. For example, Athanasopoulos and colleagues (Athanasopoulos, Damajanovic, L., Krajciová, A., & Sasaki, M., 2010a; Athanasopoulos, Dering, Wiggett, Kuipers, & Thierry, 2010b) in an extension of Thierry *et al.*'s (2009) experiment have reported perceptual shifts

modulated by length-of-stay in the new country in bilingual speakers: Greek-English bilinguals, who had been in the UK for more than 1.5 years resembled English monolingual controls more than Greek-English bilinguals resident in the UK for less than 3 months. Similar data was also reported in native speakers of Japanese (Athanasopoulos *et al.*, 2010a), albeit on the basis of behavioural measures.

Investigations in the domain of colour have provided evidence in favour of Whorfian effects in various tasks and populations (children, adults and bilinguals). The consensus is that linguistic labels affect colour cognition even at a basic visual level. Yet, a growing number of studies show that under certain circumstances these effects can vanish. For example, Winawer *et al.* (2007) as well as Gilbert *et al.* (2006) show that when participants are subjected to verbal interference (e.g., rehearsing of an 8-digit long string) the effects diminish or disappear. This begs the question of the dependent/independent nature of perceptual effects from real-time language operation. In other words, does language interference mess up selectively with language or does it have an effect beyond language processing, affecting overall cognition and even perception itself? It is in fact unlikely that a language interference task will have effects strictly limited to the language sphere, even though this has been considered a given.

3.2. Grammatical distinctions and cognition

Grammatical Number. There have also been a lot of experimental investigations of the LRH in the grammatical domain. Stepping away from basic lexical distinctions relating to basic object properties (visual or other), grammatical distinctions, which can be arbitrary and are restricted to language, have been considered an excellent test case of a pure influence of language on thought. In fact, Lucy's (1992b) revival of the hypothesis was based on a comparison of how different languages mark number on nouns and the potential effects such differences may have on non-linguistic categorisation of countable and non-countable objects. In his comparison between English and Yucatec, Lucy (1992b) observes that while English distinguishes between countable inanimate nouns (e.g., *three dogs*) and non-countable inanimate nouns, which cannot be marked for plural (e.g., **three butters*) and require a classifier (or unitiser) in order to be quantified (e.g., *three blocks of butter*), Yucatec does not allow inanimate nouns to be pluralised in any other way than with a classifier (e.g., *un-tz'it kib* "one **long thin** wax, one candle" examples and translations from Lucy, 1997). To investigate whether this grammatical difference has any impact on object categorisation, Lucy (1992b) presented participants with three sets of six pictures comprising of animals (which both Yucatec and English pluralise with a morphological marker), inanimate objects (pluralised with morphological marker in English but via a classifier in Yucatec) and non-countable substances (which both languages must pluralise externally via classifiers). In each set, the first picture represented one object and the other five presented alternate objects (of the same type) varying in number (two dogs vs.

one dog, two shovels vs. one shovel, two puddles of mud vs. one puddle of mud, and so on). Participants were asked to decide which alternate picture of each set was “most like” the original picture. The choices made by Yucatec speakers were so that they treated number changes in the animal set as most significant, whereas English speakers treated changes in animals and inanimate objects as more significant. In other words, the two choices made by the two language groups reflected the linguistic patterns of their respective language. Similar results were obtained in other classifier languages such as Japanese (Athanasopoulos, 2006). Furthermore, when Japanese-English bilingual speakers undergo a similar test, results from the bilinguals who are most advanced in English, pattern closer to those of English monolingual controls (Athanasopoulos, 2006).

Another potential consequence of this grammatical distinction lies at a semantic level. Indeed, the classifiers used to pluralise noun phrases in Yucatec (but also in Japanese or Chinese, for example) usually refer to the substance or material composition of the objects they classify, which through exposure may draw attention to such aspects in speakers of classifier languages as opposed to speakers of English (Lucy, 1992b). Lucy compared Yucatec to English speakers on a similarity judgement task where participants were presented with a target stimulus (e.g., a sheet of paper) and two alternatives, one of which was the same shape as the target (e.g., a sheet of plastic) and the other a different object made up of the same material (e.g., a book). He found that contrary to English speakers, Yucatec speakers exhibited a bias toward alternatives similar in material (e.g., the book vs. the sheet of plastic). Similar results were obtained by

Imai and Gentner (1997) when they compared Japanese- to English-speaking children. Cook, Bassetti, Kasai, Sasaki, and Takahashi (2006) extended these results to Japanese L2 speakers of English and showed that the Japanese speakers with the longest period of living in an English-speaking country showed a pattern that was between English native speakers and Japanese monolinguals; with the preference for shape (English-like pattern) increasing as a function of length-of-stay in the L2 country. Athanasopoulos (2007) sought to disentangle the influence of factors such as L2 proficiency, length of stay, and language context to test whether bilingual speakers have separate cognitive representations accessed depending on the language used in the experiment, or whether these cognitive changes are of a deeper nature, such that bilinguals may shift their cognitive patterns even when instructed in their native language. Regression analyses show that second language proficiency was the best predictor of the degree of conceptual shift towards second language categorisation preferences, independently of extra-linguistic factors such as length of stay in the L2 country and language context. A similar cognitive shift towards the L2 was also observed even in bilinguals who had never lived in the L2-speaking country before, but who had a very advanced knowledge of English (Athanasopoulos & Kasai, 2008).

Grammatical Gender. Most of the world's languages feature the notion of gender. That is, their syntax make a difference between male and female animals and humans. In those terms, gender, in those languages, can be referred to as semantic or biological gender. If gender were only semantic, studying its influence on conceptualisation would not be very interesting since semantic

gender merely reflects a distinction that is readily available in the natural world. In that sense, language reflects the world it describes but does not “dissect” it, to refer to Whorf’s (1956) quote. However, in many languages common nouns are also assigned a gender, the number of which varies between usually two and three (masculine, feminine and/or neuter) whether these nouns have a biologically relevant property or not. This marking is usually present on the noun itself as an inflection (example 1a., below) and has usually a key role in determining agreement among words in noun phrases (example 1b., below), in the selection of pronominal forms (example 1c., below), and even in subject-verb agreement (example 1d., below). Such a phenomenon is usually referred to as *formal* or *grammatical* gender since it is merely a syntactic feature independent from meaning (Aronoff, 1994; Corbett, 1991) (examples 1e.).

Consider the following examples:

1a. morphological marking of gender on the noun ‘o’ ending indicates masculine and ‘a’ feminine in Spanish.

e.g., *teléfono* ‘cat-MASCULINE’ vs. *cama* ‘bed-FEMININE’

‘telephone-MASCULINE’ vs. ‘bed-FEMININE’

1b. noun phrase agreement between noun and adjective.

e.g., *pequeño teléfono* vs. *pequeña cama*

‘small-MASCULINE telephone-MASCULINE’ vs. ‘small-FEMININE bed-FEMININE’

1c. selection of pronominal forms

e.g., *donne le moi* (masculine referent) vs. *donne la moi* (feminine referent)

‘give it-MASCULINE me’ vs. ‘give it-FEMININE me’

1d. Subject-verb agreement

e.g., *c’est la fille plutôt que le père que j’ai invitée.*

‘it is the-FEMININE girl-FEMININE rather than the-MASCULINE father-MASCULINE that I invited-FEMININE.’

1e. i. *Une sentinelle* ‘a sentry’,

ii. *Un-MASC grille-pain* (French) vs. *una-FEM tostadora* (Spanish) ‘a toaster’

iii. *Die Frau* ‘the woman’ vs. *das Mädchen* ‘the girl’

iv. *Une baleine* ‘a whale’, feminine form used for both male and female.

The examples in 1e clearly illustrate the independence of grammatical gender from a noun’s semantic content, mainly in the case of inanimate objects but sometimes even for animate and biologically gendered referents. Although the classes involved in grammatical gender are transparently related to biological gender, their use in relation to unsexuated objects is necessarily arbitrary.

Neuroimaging data suggests that grammatical gender elicits activations that appear to be independent of semantic networks (Miceli, Turriziani, Caltagirone, Capasso, Tomaiuolo, & Caramazza, 2002). The examples above also show that grammatical gender is a critical feature of language production, as it requires a number of complex operations, and therefore potentially substantial consequences for cognitive control. Finally, grammatical gender, when it is

absent in a language, cannot be replaced by other lexicalisation patterns (unlike classifiers). Grammatical gender has therefore been considered a good test-case for investigations of the LRH given its semantic transparency combined with its arbitrariness for unsexuated objects. Three kinds of approaches have been attempted:

- i. The first set of studies, which features some of the earliest attempts of investigations of the LRH, asked participants to associate gender properties to a series of stimuli. One of the first rating studies was conducted by Clarke, Losoff, Dickenson, McCracken, and Still (1981) who asked Arabic and English speakers to rate a series of words on a potency scale going from extremely masculine to extremely feminine. The results show that Arabic speakers aligned their judgements with the Arabic gender of the words significantly more than English speakers. However, in one of the stimulus sets, in the case of words whose Arabic gender was deemed “consonant” with either masculine or feminine expectations, the two groups did not differ significantly. The authors suggest that both language groups may have had expectations about gender and that the linguistic cues (in Arabic) were not sufficiently strong to produce significant differences between the two in that set. While this task made the factor under scrutiny very obvious, subsequent studies have attempted to correct this issue by asking German and Spanish speakers to rate nonwords on a potency scale to avoid culturally loaded content (Konishi, 1994). The results show that German speakers rated stimuli preceded by a masculine article (*der*) more masculine (higher on a potency scale) than stimuli associated with a feminine article (*die*). However, the Spanish speakers did not seem to be influenced by the

grammatical gender given to nonwords. Note that Konishi's (1994) paradigm, which did not use gender-transparent nominal endings, might have biased the Spanish data. Indeed, while gender-transparent endings should not be a problem for the German stimuli since German nouns are preponderantly marked for gender by the preceding article. Spanish, on the other hand has a strongly regular gender inflectional system, which in most cases requires both the article *and* the noun to carry gender information. The absence of a marking on nouns in the experiment may therefore have disengaged biological gender biases in the Spanish participants. The paper does not discuss this hypothesis, but this may mean that Konishi's task is basically a linguistic task bearing little relationship with the LRH. Overall, results from this first set of studies are mixed and inconclusive, and tend to fail at implementing experimental designs where participants could not be aware of the manipulation at hand.

ii. The second set of studies provides a series of significant improvements from the data reviewed so far by departing from potency judgements through the use of more "covert" gender assignment tasks. In a seminal study, Sera *et al.* (1994) used a voice-attribution paradigm to investigate Spanish and English speakers' object conceptualisation. Participants were presented with line drawings or words for inanimate objects, half of which were natural objects (e.g., tomato, fire, eye) and the other half was composed of artificial objects (man-made; e.g., hat, knife, helicopter). These two categories were subdivided in two with one half of the objects being masculine in Spanish and the other half feminine. Participants were asked to imagine whether each of the objects would speak in a male or female voice if it were to be featured in a cartoon. The results

show that Spanish speakers tend to assign a voice congruent with Spanish gender, while English speakers tend to assign voices in line with natural/artificial distinctions (more female voices for the natural items and more male voices for the artificial items). Although Spanish speakers were more influenced by item gender, subsequent analyses show that they also exhibited a trend towards the natural/artificial distinction. Sera *et al.* (1994) later implemented their voice-attribution paradigm with children from 5 to 10 years of age in an attempt to disentangle whether Spanish children start with an English-like pattern of categorisation and then move on to gender-influenced categorisation. Spanish children start to show grammatical gender influence from the age of 7. In a subsequent study Sera *et al.* (2002) obtained similar effects with French-speaking children and adults but failed to replicate the effect with German speakers. Bassetti (2007) replicated the results of Sera and colleagues (Sera *et al.*, 1994; 2002) in a slightly more controlled online version of the voice attribution task with monolingual Italian children. However, her group of Italian-German bilinguals did not show any reliable influence of gender, potentially because of the three-way gender assignment system in German. Kurinsky and Sera (2010), further showed that English native speakers, who learn Spanish at an adult age as a second language in a university classroom setting, start to be influenced by Spanish gender on a voice attribution task already after 10 weeks of instruction. However, in this study the authors repeated the same stimuli across testing sessions, which may have caused the effect. Replication data collected by Boutonnet and Athanasopoulos (unpublished manuscript) show that Kurinsky and Sera's (2010) effect may have reflected stimuli repetition. Although the

results from these studies are promising, voice attribution paradigms make the reference to gender too obvious and therefore these paradigms are very susceptible to participants' strategy in resolving the task.

Influences of grammatical gender in German speakers have, however, been revealed in a series of studies conducted by Boroditsky and colleagues (Boroditsky *et al.*, 2003; Boroditsky & Schmidt, 2000; Phillips & Boroditsky, 2003). Boroditsky and Schmidt (2000), in an all-in-English memory task, taught German-English and Spanish-English speakers object name- (e.g., chair) person name (e.g., Mary) pairs. All items were selected so that they would have opposite genders in German and Spanish. Participants were asked to recall the gender of the proper name that had been associated with each object name. Both groups were significantly more accurate when the genders of the proper names coincided with those of the objects. For example, German speakers were better at remembering *apple* (masculine in German) when it was associated with *Patrick* than when it was associated with *Patricia*. The Spanish speakers produced the reverse pattern of results for the item 'apple' (since it is feminine in Spanish). However, the same criticism as for voice attribution paradigms can be made of this study, albeit to a lesser extent. Indeed, it is likely that participants were cued to pay attention to gender since the person names explicitly contrasted on a gender basis.

iii. Finally, whilst remaining in the domain of behavioural measurement, the third set of studies have pushed further in minimising participants' awareness of gender manipulations. Martinez and Shatz (1996) investigated the strategies Spanish- and English-speaking children use to classify objects in a free-sorting

categorisation task. Participants were given 10 line drawings of four items with natural gender (2 males, 2 females) and 6 drawings of inanimate objects of which three had a masculine gender and the other three a feminine gender in Spanish. Children were asked to put together the drawings that “go together”. Their results suggest that the Spanish children are more likely to be influenced by the grammatical gender of the object names compared to matched English speakers who were influenced by the artificial/natural nature of objects.

Vigliocco, Vinson, Paganelli, and Dworzynski (2005) also designed a study which hid all references to gender by testing Italian, German and English speakers on a similarity judgement task based on semantic meaning. Participants were presented with triplets of words, which referred to animals and artefacts and were asked to choose which of the images (contrasting in genders) was more related in meaning to the target. Judgements from the Italian speakers suggested that they were influenced by Italian gender, although this was only the case for words depicting animals. German speakers tested in a similar task did not even exhibit the effect for animal words. Again here, the gender effect tended to disappear in German, which given its three-way gender system may lead to more inconsistent mappings. A subsequent study by Ramos and Roberson (2010), provided somewhat contradictory results since they managed to obtain an effect of grammatical gender on both animate and inanimate objects in Portuguese speakers using a similar triadic judgement task. However, the effect disappeared with pictorial stimuli. Ramos and Roberson (2010) also conducted a replication of Sera *et al.*'s (1994) paradigm suggesting that for gender effects to arise, grammatical gender must either be relevant to

the task (as in the case of voice attribution) or participants must be engaged in linguistic processing.

In a different vein, Boroditsky *et al.* (2003) refer to a study in which they collected free adjective elicitation data from speakers of Spanish and German who were asked to produce the first three adjectives that came to their minds in relation to pictorial stimuli. The adjectives produced by the participants were then rated by a series of other participants on a masculine to feminine potency scale. Results showed that Spanish and German speakers generated adjectives rated as more masculine for items whose names had a masculine gender and the opposite for items with a feminine gender. Since item sets were selected to be of opposite genders in the two languages, the groups showed reverse patterns. This means that when participants saw a *bridge*, which is feminine in German, but masculine in Spanish, the German group generated adjectives such as *beautiful*, *elegant* and *peaceful* while the Spanish speakers described it as *big*, *strong* and *towering*. With an item such as a *key* (feminine in Spanish, but masculine in German), Spanish speakers produced adjectives such as *little*, *lovely* and *tiny*, whereas German speakers described it as *hard*, *heavy* and *jagged*. From this data, it almost seems like both speakers group conceptualised the objects as an old and big key or a well designed, modern and aerodynamic bridge in the case of the German speakers whereas the Spanish group considered the key to be small and cute, and the bridge to be big and rusty.

One of the most recent, and perhaps most convincing evidence of grammatical gender-related relativity comes from works by Cubelli and colleagues, in which the authors minimised grammatical gender awareness. In a

first experiment, Cubelli, Lotto, Paolieri, Girelli and Job (2005) asked Italian participants to name a picture while ignoring a distractor word presented at the same time (picture-word interference). Participants were slower at naming when picture and distractor words had the same grammatical gender. The cumulative interference effect shown here suggests that grammatical gender properties are selected whenever a noun has to be produced and this even occurs outside of a sentential context. Cubelli, Paolieri, Lotto, & Job's (2011) study set out to investigate this possibility on a non-linguistic level, therefore better tackling the question of linguistic relativity. In this study, participants were shown pairs of pictures, which were either semantically related or not and were asked to judge whether the two pictures belonged to the same semantic category via button presses measuring participants' RTs. Furthermore, the participants were unaware that half of the semantically related and unrelated pairs were gender congruent and the other half were incongruent. A semantic priming effect was found (participants were faster at making decisions when the two pictures were related than unrelated) and their RTs were also modulated by gender congruency such that pairs of pictures of which the name had the same gender were categorised faster, irrespective of semantic congruency. In order to avoid potential bias in the picture materials, Cubelli *et al.* (2011) conducted a follow up experiment comparing Italian and Spanish speakers on the same task with the stimuli pairs selected to be gender congruent in one language but incongruent in the other. As predicted, the gender congruency effect was reversed in the two languages such that a masculine congruent pair in Italian produced faster RTs whilst the same pair produced significantly slower RTs in

the Spanish participants. This follow-up experiment, therefore, provides strong evidence for a gender effect on object categorisation since the design was fully counter-balanced and each of the two participant groups served as control group for the other.

Overall, investigations of the LRH in the grammatical domains of number and gender have shown that grammatical features of languages, which require speakers of different languages to focus on some aspects of reality more than others when they produce or understand words and sentences, may have their conception of objects altered even in non-linguistic contexts. Furthermore, Whorfian effects prompted by features like number and gender seem to arise from an early age. Studies in children in both domains have reported evidence that categorisation strategies or biases are altered from the moment such features occur in speech production. Effects of language on object categorisation also appear to be dynamic, as has been shown in bilingual and second language speakers, where the effects of the native language can be suppressed or altered by second language acquisition, even in late learners and even when L2 learning is experienced without immersion in the L2 environment (Athanasopoulos, 2006; 2007; Athanasopoulos & Kasai, 2008; Kurinski & Sera, 2010).

In contrast, a wealth of studies have failed to establish any significant effect of language on non-linguistic cognitive operations. For instance, Kousta, Vinson, and Vigliocco (2008) have strongly argued against an effect of grammatical gender on non-linguistic, conceptual representations. As already mentioned above, Vigliocco *et al.* (2005) have stressed the limited effect of language on

conceptual representations to the domain of animate stimuli, when gender is relevant (i.e., semantically motivated). Kousta *et al.* (2008) used a continuous naming task asking Italian and English monolinguals as well as Italian-English bilinguals to name pictures presented at a fast rate. This type of paradigm is known to produce semantic substitution errors (e.g., saying *eye* when *ear* is intended) as a result of competition between semantically related candidates. Their hypothesis follows the logic that if grammatical gender alters speakers' conceptual representations, semantic similarity between gender sharing competitors should be increased and Italian speakers should produce more gender-preserving errors than English speakers. Kousta *et al.* (2008) found that Italian monolingual speakers produced a significantly higher number of gender-preserving substitutions than the English monolinguals. However, Italian-English bilinguals performed exactly like monolingual speakers in either of their languages depending of the language of the task. The authors take this as evidence that conceptual representations, which they consider language-independent, are not modified by knowledge of gender and that the influence of language is limited to what they call 'language-specific thought' (i.e., a weak account of the LRH). Finally, similar to some studies in the colour domain, when participants are placed under verbal interference conditions, the effect of gender on categorisation has been shown to disappear. In one of their experiments, for instance, Cubelli *et al.* (2011) showed that the gender priming effect mentioned earlier disappears when participants are asked to continuously repeat "bla, bla, bla" while performing the semantic task. While the effect of semantic relatedness is preserved, the gender effect disappears. Such results

weaken the effect of gender on object categorisation and lead the authors to argue, like Kousta *et al.*, in favour of a weak, limited version of the LRH.

3.3. Lexical items, object perception and categorisation

Much in the same fashion as colour studies, the effect of lexical labels has also been investigated in relation to higher visual processing, such as object categorisation. Indeed, if language-specific terminology has an effect on the perception of stimulus characteristics (such as colour), it should extend to the perception of more complex common objects. Lupyan, Thompson-Schill, & Swingley (2010) investigated the modulation of physical perception using simple letter stimuli. 'B-b' and 'B-p' letter pairs have equal visual similarity (b, being the symmetrical reverse of p and *vice versa*). However, conceptually B-b pairs are more related by virtue of having the same label, whereas the B-p pair involves two distinct labels. Participants were presented those pairs, and asked to perform speeded same-different judgements. The authors report a category-effect whereby participants' RTs were slower on within-category compared to between-category pairs. These effects, however, disappeared when the letters were rotated at a 90° angle, which is supposed to have reduced stimuli familiarity and prevented those stimuli to be identified as the letters *b* and *p* (cf. **fig. 2**, for example of the stimuli).

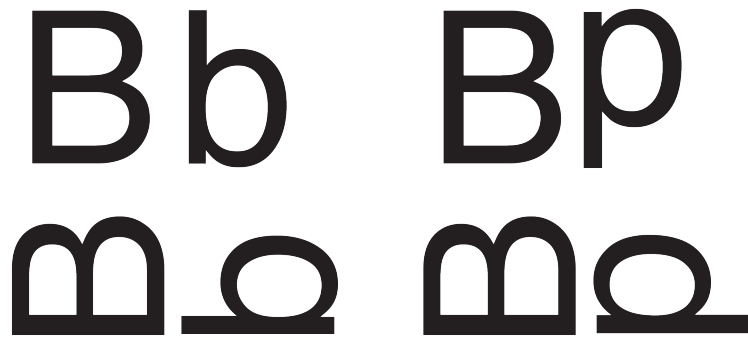


Figure 2. Illustration of Lupyan, Thompson-Schill, & Swingle’s (2010) experimental stimuli. Right part of the figure depicts upright and 90° rotated within-category stimuli pairs (B-b). Left part of the figure depicts upright and 90° rotated between-category stimuli pairs (B-p).

Similar results were obtained in an earlier study (Lupyan & Spivey, 2008) using distorted versions of the numerals 2 and 5 (cf., **fig. 3**) in a visual search paradigm where the stimuli could be taken as abstract symbols or perceived (after specific instructions) as rotated 2s and 5s. Participants who were told to consider the stimuli as rotated numbers were faster in the visual search task than those who had no instruction.



Figure 3. Illustration of Lupyan & Spivey (2008) experimental stimuli representing the symbols which could be taken as rotated distorted versions of the numerals 2 and 5.

In a straightforward attempt to extend the evidence of Whorfian effects beyond the realm of colour, Gilbert *et al.* (2008) adapted a version of the lateralised visual search task employed by the same authors (Gilbert *et al.*, 2006) from

colour stimuli to silhouettes of two cats and two dogs. Stimuli were arranged so that a picture of a cat among an array of dogs (and *vice versa*) made up a between-category condition, by virtue of the two stimuli being referred to by a different label. The within-category trials featured a picture of a cat among an array of other cats (and *vice versa*), where the two stimuli had the same label. The results show that labels modulated participants' performance, whereby faster detection rates were obtained in between- than within-category trials. Additionally, the authors reported stronger effects when items were presented in the right-visual field (processed mainly by the left hemisphere of the human brain) as compared to the left-visual field. This was taken as evidence that linguistic labels affect object perception. It is important to note, however, that basic perceptual differences between stimuli were not controlled. The fact that between-category trials lead to faster RTs could at least in part be due to the fact that the stimuli are simply more distinguishable in a visual search task given that cats and dogs are perceptually more different than two cats or two dogs. Without data from a control group with a different linguistic background or a perceptual baseline it is impossible to rule out low-level perceptual effects even in light of the visual field interaction reported (or the additional split-brain patient data offered).

Overall, the results reviewed support a far-reaching influence of language on categorical and perceptual processes. However, the same limitations as those outlined above apply here, since participants exposed to verbal interference see Whorfian effects substantially drop in power (Gilbert *et al.*, 2008).

3.4. Strong effects that are weak, a problem?

One could bluntly answer ‘no’ to this question and would probably be correct. Indeed, the selective review of the literature above shows the diversity of the approaches and backgrounds characterising researchers in the field and the discrepancies in terms of methodological approaches must account for some of the inconsistencies in the conclusions drawn.

However, answering “no” bluntly could be taken as ignoring the elephant in the room (a critique originally made by Lupyan [2012] as well as Wolff & Holmes [2011]). A general consensus has emerged in the literature, that for Whorfian effects to be of any scientific interest, they have to manifest themselves in non-linguistic cognition, that is, independently of language processing and under any circumstance. Also, preferably, these effects should “run deep” and be hard to disrupt. Although it is very unlikely that Whorfian effects are deterministic as pointed out earlier, sceptics often stress this impossibility as an argument against linguistic relativity. Results showing that language needs to be at the forefront of a task in order to affect participants’ performance (Slobin, 1996; 2003) have led to limiting interpretations of the LRH (cf. “*thinking for speaking*” [Slobin, 1996]) and these effects have often been regarded as uninteresting or unsurprising, which falls in the realm of critiques such as Pullum’s (1989) gritty reaction to observations about the Eskimo language reported by Boas:

“Even if there *were* a large number of [words] for different snow types in some Arctic language ... this would not, objectively, be intellectually interesting; it would be the most mundane and unremarkable fact.” (pp. 278–279)

Even more damaging, or perceived as such, is the flurry of evidence showing that under conditions of verbal interference the “strongest” Whorfian effects diminish or disappear all together (e.g., Winawer *et al.* (2007), Gilbert *et al.* (2006), Drivonikou *et al.* (2007), for colour; Kousta *et al.* (2008), Cubelli *et al.* (2011), for grammatical gender and Gilbert *et al.* (2008) for objects’ lexical labels). Most researchers faced with the quasi-deterministic take on Whorfian effects have been forced to put a damper on their claims and have fallen short of convincing and intellectually interesting accounts. This has led to the idea that effects of language on processes such as categorisation or basic visual processing can be removed easily and must be superficial. In other words, although no one would be ready to accept a deterministic influence of language on thought, it is precisely what everyone seems to want established in order to be convinced. Such a lack of theoretical unity has led to branchings of the LRH as reflected in a recent review by Wolff and Holmes (2011). Indeed, the literature seems to distinguish five different strands. At the weaker end of the spectrum is Slobin’s (1996) conceptualisation where thinking is affected *before* language production. Other interpretations propose that we think *with* language such that language is seen as a meddler, where linguistic representations compete with non-linguistic representations or that language is an augments, where linguistic representations extend or enable non-linguistic representations. The last group of interpretations contends that thought is affected *after* language. In that case, language is seen as a spotlight, where it makes certain properties of objects (concrete or abstract) more salient than others, that is, it is considered as an inducer, where it primes certain processes in non-linguistic space. The lack of

conceptual unity in this regard is clearly reflected in those different branches which individually account for some, but never all, of the data.

However, is the “strong/weak” paradox real or is it not just a consequence of ill-driven and obsolete theoretical accounts? In fact, the very formulation of the phrase “influence of language *on* thought” is misleading since it presupposes a clear separation between, on the one side, linguistic representations and on the other, non-linguistic representations. Therefore, the origin of the problem is the commonly accepted, and hardly ever questioned, assumption that there are several kinds of processing: verbal, language specific, independent and modular (for an extreme view see, Fodor, 1975) vs. non-verbal and, most importantly, that these two do not really interact and are not part of the same system. Such a conception seems to commit accounts of Whorfian effects to the pitfalls of determinism and results in the multiplication of accounts.

Fortunately, the last few years have seen the emergence of alternative and novel accounts of Whorfian effects (Lupyan, 2012), which have adopted modern psychological models of cognition. Furthermore, recent neurophysiological evidence (some of which reviewed in section 1.2 above) on the physiological underpinnings of categorisation lends a picture of cognition which is highly distributed, interconnected, and interactive. The view we adopt in the thesis is in line with distributed accounts of human cognition, i.e., the label-feedback hypothesis (Lupyan, 2012). We will see that this research programme allows researchers to formulate clear and testable hypotheses as well as explain their results in a much more psychologically integrated view of human cognition.

4. Interactive account of Whorfian effects: *the label-feedback hypothesis*

4.1. Online and transient effects of labels on perceptual representations.

The label-feedback hypothesis is based on the fact that linguistic labels (names) are key elements for categorisation and cognition in humans. We have outlined this idea earlier in the thesis by considering Harnad's (2005) proposal about "linguistic theft" (cf. **Section 1.1 above**) as well as evidence from studies of developmental category learning (cf. **Section 2.1 above**). The label-feedback hypothesis proposes that Whorfian effects are driven by transient (on-line and temporary) modulations of on-going processes such as perception or categorisation, which can lead to early effects, i.e., the kind of effects most researchers refer to as "strong". For example, the hypothesis accommodates data such as that from Thierry *et al.* (2009), showing early and robust influences of colour terminology as indexed by modulations of N1 and even P1 peaks of brain potentials. The way in which labels "wrap" perception is by the co-activation of linguistic and perceptual features, which give rise to a "hybrid visuo-linguistic experience" (Lupyan, 2012, p. 4). In such a dynamic model, it is expected that the strength of the modulation can be made stronger (e.g., when labels are explicitly activated, Lupyan & Spivey, 2010a) or weaker by disrupting access to labels under conditions of verbal interference (e.g., Cubelli *et al.*, 2011; Winawer *et al.*, 2007), or by reducing stimulus familiarity (e.g., Lupyan *et al.*, 2010; Lupyan & Spivey, 2008). An important point is that this modulation, if transient, will be expected to decay overtime after label activation. This is exactly what Lupyan & Spivey (2010a) demonstrated. Indeed, if the time

between the presentation of the label exceeds circa. 1600 ms, the effect of label on object detection vanishes.¹

4.2. Computer model simulation of online label effects on perceptual representations.

The predictions of the label-feedback hypothesis have been implemented into a simple connectionist model. Adapted from Rumelhart *et al.*'s (1986) fully recurrent neural network, the model is made of three layers. A “perceptual level” which can receive input from the outside world was connected bi-directionally (so as to allow for both bottom-up and top-down feedback) to an intermediate layer, which can be thought of as the layer that develops “conceptual” representations. The conceptual layer is in turn connected bi-directionally to a label layer, containing categorical labels, as well as back to itself. The model was trained on two abstract categories of stimuli {for the detailed methodology see, Lupyan:2012fb} to learn and produce names, given stimuli, and to return the properties of learned stimuli upon “hearing” labels corresponding to the stimuli. Following training, access to the label layer can be turned *on* or *off* or can be provided exogenously along with stimuli presentation. In the first case (meant to represent the default mode of processing), the access to the label layer is *on*, the label is part of the network and, upon stimulus presentation, it can modulate perceptual processing via top-down feedback. When tested in this condition the network successfully differentiates two categories with next to no overlap. In a condition where

¹ Note that although we refer to experiments with explicit references to labels, the hypothesis does predict spontaneous label activation (e.g., Gilbert *et al.*, 2008; Lupyan & Spivey, 2008). The introduction of labels in those experiments is for the purpose of testing the extreme boundaries of such label effects.

access to the label layer is not allowed to feedback, the network reverts to distinguishing stimuli on a perceptual basis only. Finally, when labels are provided externally with the stimuli, categorisation success is further enhanced. These three test-cases replicate with high fidelity the outcomes reviewed in the experimental literature: both strong and weak effects with only one process and different levels of access to labels.

4.3. Neural underpinnings of the hypothesis.

At first, the idea that complex (high-level) mental representations can influence low-level and rapid perceptual processes may seem preposterous if one adopts the pervading view in linguistics and psycholinguistics that the brain is essentially a feed-forward system. However, this conception has recently received substantial criticism. Indeed, a growing body of research in the past 20 years have provided evidence and argued in favour of a more refined conception of the brain as a not-so-hierarchical system (Churchland, Ramachandran, & Sejnowski, 1994; Foxe & Simpson, 2002; Freeman, 2000; Gilbert & Sigman, 2007; Koivisto, Railo, Revonsuo, Vanni, & Salminen-Vaparanta, 2011; Lamme & Roelfsema, 2000; Mesulam, 1998). Lamme & Roelfsema (2000), for example, review data showing that prefrontal areas can prepare the visual cortex to perceive particular dimensions of stimuli *before* they are actually displayed. Effects of verbal labels in the label-feedback hypothesis would therefore rest on the highly interactive organisation of the human brain, whereby the processing of an object name triggers a series of feedback activations to object-selective areas serving as a predictive signal to the visual system.

4.4. Implication for LRH research

The label-feedback hypothesis has so far been tested mainly in the domain of the relationships between lexical labels and object perception (see, Lupyan:2012 for review) and makes little predictions regarding the potential influence of grammatical features on the perception and categorisation of objects. Therefore, more investigations in this area are needed. Additionally, what becomes clear from this account of Whorfian effects is that researchers need to tackle the online time course of such effects, using methodologies that can inform their neural sources. This is precisely what I endeavoured to do in this thesis, which presents data from a combination of behavioural and neurophysiological experiments. In order to look at the unfolding in time of Whorfian effects, as well as shed some light on their neural basis, the experiments presented here have used event-related potentials.

Chapter 3

Event-Related Potentials in language
and perception

Electroencephalography (EEG), as a neuroimaging technique, has a record for having been developed and used for the longest time. The first human EEG was recorded by Hans Berger in 1924, and has since been used for a large array of purposes. Clinically, it is commonly used in diagnosing epileptic seizures, catatonia or determining whether a person is alive or dead. Nowadays, EEG is also widely used for research in a large number of fields in neuroscience.

All of the experimental work presented in this thesis is based on behavioural and Event-Related Potentials (ERPs) measures. Here, we offer an overview of the general methodological aspects of the acquisition and analysis of ERP data, their advantages and their limitations. This section is followed by a description of the biological underpinnings of the EEG signal. Finally, we present the relevant applications of ERPs to the domain of language processing and linguistic relativity.

1. Event-Related Potentials, a particular analysis of EEG signal

1.1. EEG recording and setup

EEG signal is obtained by recording the electrical activity produced by the brain, which is recorded with the help of electrodes set on different scalp regions. The signal at any electrode site is obtained from the difference in electrical potential between this electrode and a ground electrode relative to the difference between a reference electrode (mastoids, or common average) and the ground. EEG systems record electrical activity from any number of sites up to 256 electrodes evenly distributed over the scalp according to the 10/20

In the setups used to conduct the present studies, electrodes are not in direct contact with the scalp, since electrodes are placed on a plastic mount in an elastic cap and due to the presence of hair or any other barriers. Care must therefore be taken to reduce electrical impedance through the application of a conductive and mildly abrasive gel, which establishes a bridge between the scalp and the electrodes and maintains resistance below $5\text{k}\Omega$ for the scalp electrodes and below $10\text{k}\Omega$ for external electrodes. Due to its weakness, the electrical signal produced at the surface of the scalp must be amplified close to the electrodes because it decays rapidly in electrical wires which have their own impedance. Signal is sampled at 1 kHz , representing 1 point of data per millisecond.

1.2. From EEG to ERPs

Most commonly in neuroscientific research, EEG data is used to derive ERPs. While EEG is a simple recording of spontaneously occurring brain potentials, ERPs are derived in the context of repeated stimulations (events) which, when averaged over trials reflect the potentials related to the stimulation only since random and unrelated potentials (noise) are “cancelled out” given the fact that noise is not systematically phase-locked to the presentation of the stimuli.

Following an offline filtering of the continuous EEG data through a zero phase shift digital filter set at 30 Hz (48 dB slope) designed to reduce contamination from exogenous electrical noise (e.g., 50 Hz electrical devices present in and around the testing room), the continuous data is mathematically

corrected for artefacts caused by vertical eye movements (blinks) using a method provided by Gratton, Coles & Donchin (1983), and periods where the activity exceeds $\pm 75 \mu\text{V}$ are automatically rejected. The entire EEG recording is usually re-referenced to either a common average reference or a bi-mastoidal reference (consisting in the average of the activity recorded from the electrodes placed on the mastoids TP9 and TP10 on **fig. 1**). Re-referencing is essential to (1) prevent distortion of the data by localised electrical artefacts and (2) to preserve the data from spurious topographical asymmetries. Finally, the continuous data is segmented into time windows of usually 1000 ms following and time-locked to stimulus onset, which are averaged across trials of similar stimuli type. These individual averages are generally averaged across participants (grand average) so as to cancel individual differences, which are considered meaningless or uninteresting, at least in the research presented in this thesis.² Mean amplitudes measured around peaks of interests (rather than peak amplitudes) are then used to quantify the observed effects and are generally subjected to factorial Analyses of Variance (ANOVAs).

Note that the method outlined in this section is one of many ways to obtain ERPs. Different references, different criteria for artefact rejections and different methods for eye-blink corrections are known to lead to differences in the derived ERP. Although the method above is the one preferred in my lab, I have always made sure to compute different averages and artefact rejections in order to assess whether the effects still hold under different practices.

² Individual differences can of course be analyzed (if such a variable is of interest) with methods that are less common and fairly novel such as single trial analyses.

1.3. Waves with meaning: ERP components

As explained earlier, the assumption behind the analysis of ERPs is that, by way of averaging over a certain number of trials, only the activity that is recurrent and time-locked with stimuli presentation will survive the process. The resulting output is characterised by a series of positive and negative peaks, referred to as components and labelled according to their polarity (P for positive, N for negative) and their approximate latency (e.g., P100 for the first positive wave appearing around 100 ms, cf., **fig. 2**).

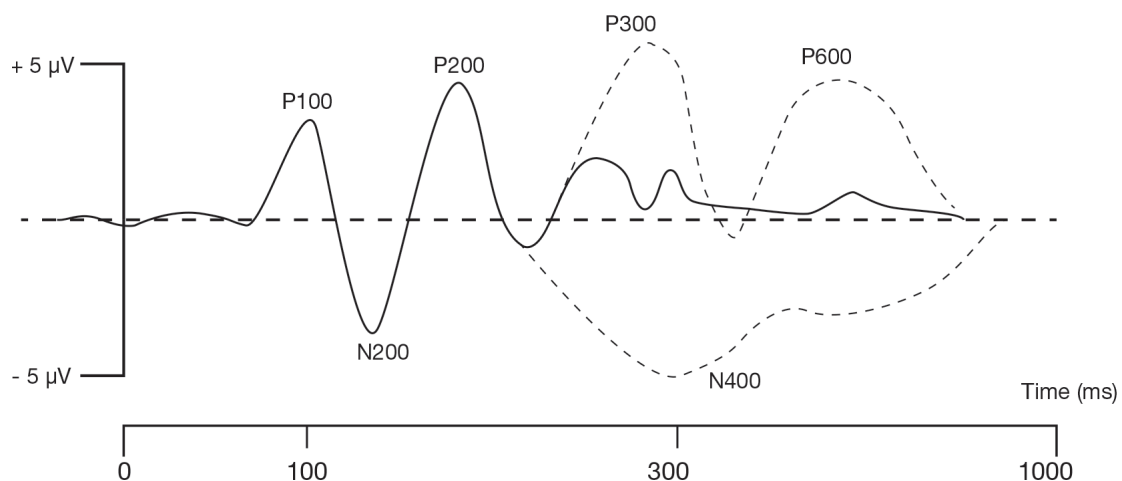


Figure 2. Illustration of some of the most studied ERP components. Solid lines depict basic perceptual components present in any visual experiment. Dotted lines represent “optional” components, which may be elicited depending on the experimental paradigm.

The early peaks (P100, N100, P200) are generally associated with basic low-level perception and are often considered to reflect automatic processes. That is to say, these components should be elicited as long as something perceptual (word,

picture or sound) is presented to the participant. The later, more conscious, components may or may not be elicited since they depend on the experimental context. Although the generalisation we just made is valid in most cases, it should be noted that the functional significance of each peaks is often task-dependant and therefore, comparison of early vs. late or automatic vs. conscious should not be made without taking into consideration the task at hand. A review of the main ERP components as well as their elicitation and interpretation will be given in the next section.

1.4. The neural source of EEG signal

The brain produces two types of electrical activity: action potentials and postsynaptic potentials. Action potentials are discrete electrical spikes, which travel from the beginning of the axon at the cell body to the axon terminals, where neurotransmitters are released. Postsynaptic potentials are voltages that are produced when the neurotransmitters bind to the receptors on the membrane of the postsynaptic cell, causing ion channels to open or close. While both types of potentials can be recorded, EEG can only reliably record postsynaptic potentials.

Indeed, EEG electrodes placed at the surface of the scalp cannot detect action potentials due to their timing of the potentials and the physical arrangement of the axons. Action potential are generally brief and very localised voltage spikes. Postsynaptic potentials on the other hand last longer than action potentials (tens or even hundreds of milliseconds). Additionally, postsynaptic potentials

occur essentially instantaneously rather than travelling down the axon at a fixed rate, which allows these potentials to summate and make recording at the scalp possible.

Little is known of the actual biological events that give rise to ERPs but the best and current estimation contends that if an excitatory neurotransmitter is released at the apical dendrites of a pyramid cell, current will flow from the extracellular space into the cell, yielding a negativity on the outside of the cell in the region of the apical dendrite. Current will then also flow out of the cell body and basal dendrites, yielding a positivity in this area. Such a process leads to the creation of a *dipole*, which if limited to a single neuron would be too small to be recorded from the surface of the scalp. However, under the right conditions, the dipoles from many spatially aligned neurons (thousands or millions), as is mostly the case for pyramidal cells, which are aligned perpendicularly to the surface of the cortex, will produce a potential at roughly the same time. This will allow for the summation of postsynaptic potentials and make recording at a long distance possible (for an exhaustive account see, Luck, 2005).

1.5. Advantages and limitations of ERP data

In order to study human cognition, EEG comes with the great advantage that it is a completely non-invasive technique; it is silent and relatively cheap compared to other more “heavy” neuroimaging techniques such as functional Magnetic Resonance Imaging (fMRI), Positron Emission Tomography (PET) or Magnetoencephalography (MEG). EEG is fully compatible with other

neuroimaging methods as well as brain stimulation techniques like tDCS and TMS and is nowadays used greatly in conjunction with those. Most importantly, EEG offers the best resolution in terms of time with smaller than a millisecond precision. Considering that a good estimation of the speed at which neurons fire is the order of 5 milliseconds, it is safe to assume that even the smallest periods of activity can be recorded.

One of the main shortcomings of EEG data, however, is its poor spatial resolution due to the fact that only a limited number of electrodes (maximum 256) can be placed on the scalp and, as explained above, only potentials from generators close to the scalp and arranged in a parallel fashion can be reliably recorded. Additionally, electrical potentials have to go through a series of barriers (essentially the skull), which result in a great deal of diffusion therefore blurring the source of the signal greatly. Consequently, classical use of EEG cannot reliably inform of the location of the measured voltages and the type of localisation usually available from fMRI cannot be replicated.

It should be noted, however, that advances in statistical modelling have enabled more trustable analysis and localisation of ERP sources as well as analyses of brain connectivity (phase-locking, granger causality etc.) with the added benefit of time usually lost in fMRI data. These methods are still in their early stages of development and some care should be taken when inferring from those data (see chapter 7 in Luck, 2005). The next section focuses on the classical uses of ERPs in the domain of language processing and will provide a selective review of the relevant components for the studies carried out in this thesis.

2. Overview of the use of ERPs in language processing and linguistic relativity

ERPs have been extensively applied to language processing and components sensitive to most domains of language (phonology, syntax and semantics) have been reliably identified. We expected a series of language related components that would be elicited if language were to influence processing on non-verbal tasks.

2.1. Meaning, relatedness and expectancy: N400

The N400 is by far one of the most studied and replicated language-related ERP component, which is probably due to its extreme robustness and reliability. The N400 was first reported by Kutas & Hillyard (1980). It is characterised by a large wave of negative polarity usually recorded from centroparietal scalp sites typically peaking in the 350 – 500 ms window. In their 1980 paper, Kutas & Hillyard elicited an N400 when presenting participants with sentences comprising endings that were more or less expected semantically. Where a sentence like “it was his first day at work” elicited very little N400 negativity, one like “he spread the warm bread with socks” elicited a strong negativity given its unpredicted and almost impossible ending. N400 modulations are not restricted to the visual modality and sentence processing since they can be elicited by single words (auditory or visual modality) as is the case in classical priming paradigms (Bentin, Kutas, & Hillyard, 1993; Bentin, McCarthy, & Wood, 1985; Holcomb & Neville, 1990), they can also be elicited by non-linguistic stimuli such as pictures or even across modality between words and pictures (Ganis,

Kutas, & Sereno, 1996; Holcomb & McPherson, 1994). The N400 has also been elicited in a flurry of domains such as maths, gestures and actions (see, Kutas & Federmeier, 2011 for an extensive review). A great advantage of this component resides in the fact that its amplitude can offer one of the most fine-tuned measures of semantic integration as it is sensitive to even the smallest differences in degrees of relatedness. The N400 therefore appears to be a solid component with which to study processes linked to stimuli integration.

2.2. Grammar and syntax processing: P600 & LAN

Syntactic violations also elicit distinctive ERP components. The P600 was first reported by Osterhout & Holcomb (1992) where participants undergoing ERP recording had to read syntactically correct or incorrect sentences. Consider the following sentences:

- a) The broker persuaded the man to sell the stock.
- b) *The broker persuaded to sell the stock was sent to jail.

Their data shows that when participants read a sentence such as (b) a bigger P600 wave was elicited. Osterhout & Holcomb concluded that the P600 indexed detection of syntactic violations and these results were consistently replicated (e.g., Friederici, Steinhauer, & Frisch, 1999; Hahne & Friederici, 2002; 1999). Recently, however, the P600 has also been associated with monitoring processes and re-evaluation (Kolk, Chwilla, van Herten, & Oor, 2003; Kolk & Chwilla, 2007; van de Meerendonk, Kolk, Vissers, & Chwilla, 2010). It therefore seems

reasonable to consider P600 to reflect re-evaluation or double-checking, which may be expected in the processing of syntactically abnormal sentences.

Syntactic violations can often be measured much earlier by the LAN (Left-Anterior Negativity) component, usually peaking in the 300–500 ms window on the frontal part of the left side of the scalp. Morpho-syntactic violations (word inflection) have been reported to elicit LANs consistently in the domain of verbal inflection and subject verb agreement (Coulson, King, & Kutas, 1998; Friederici, Pfeifer, & Hahne, 1993; Vos, Gunter, Kolk, & Mulder, 2001) as well as grammatical gender (Gunter, Friederici, & Schriefers, 2000; Thierry, Cardebat, & Démonet, 2003), which is particularly relevant for the study presented in Chapter 4.

In an attempt to unify these two components, Friederici (1995) proposed a model featuring early vs. late sensitivities to syntactic violations. In her model, Friederici distinguishes two early processes, one strictly related to word and grammatical classes, which is reflected by early modulations of the LAN, often named ELAN. This early process has however very recently been criticised (see, Steinhauer & Drury, 2012). The second early process is reflected by the LAN proper and it is supposed to reflect the processing of morpho-syntactic properties. The third and late phase, P600 modulations, is supposed to reflect the *post-hoc* integration of different streams of information and the repair of anomalies involving sentence structure, and or semantic incongruencies. It is therefore not uncommon in the literature to find that the LAN (and ELAN, provided they exist) is followed by a P600 – usually depending on tasks demands.

2.3. Similarity, deviancy and language: Deviant-Related Negativities

Deviant Related Negativities (DRN) have originally been elicited in the auditory modality in the context of oddball paradigms which present participants with a repetitive chain of identical stimuli (standards) interrupted by infrequent mismatching stimuli (deviants). DRNs are usually split in two types Mismatch Negativity (MMN) and visual Mismatch Negativity (vMMN). The MMN coincides with the N1 peak but is particular of oddball paradigms and should not be interpreted like a classical N1 functionally. The MMN has been used to study the perception of speech sounds, e.g., phonological distinctions (Näätänen, 2001) but is not restricted to linguistic sounds and can be elicited by differences in pitch, and rhythm, for example (Näätänen, 1995; 2009; Winkler & Czigler, 2012). A particularly interesting advantage of this component resides in the fact that it can be elicited even when participants are not actively paying attention to the stimuli and it is therefore thought to index automatic and preattentive processes of perception. Its amplitude is modulated by the degree of perceived similarity between the standards and deviants so that the greater the difference between the stimuli, the more negative the amplitude of the MMN is.

It is believed that there exists a visual counterpart of the MMN (vMMN), which reflects similar automatic processes in the visual modality (for a review see Pazo-Alvarez, Cadaveira, & Amenedo, 2003). The vMMN peaks at the same time as its auditory counterpart and can also be elicited by stimuli presented outside the focus of attention (parafoveal presentation, for example). The

vMMN has been used extensively to study questions linked to the influence of language on perception of colour, for example (Athanasopoulos, *et al.*, 2010b; Clifford *et al.*, 2010; Liu *et al.*, 2010b; Thierry *et al.*, 2009). We offer a more detailed review of this particular application to the domain of object perception in Chapter 5.

Chapter 4

Unconscious effects of grammatical gender
during object categorisation

This paper is published as: Boutonnet, B., Athanasopoulos, P., & Thierry, G. (2012). Unconscious effects of grammatical gender during object categorisation. *Brain Research, 1479(C)*, 72–79. doi:10.1016/j.brainres.2012.08.044

Abstract

Does language shape thought? Here, we approach this question from the perspective of grammatical gender in bilinguals. We tested Spanish-English bilinguals and control native speakers of English in a semantic categorisation task on triplets of pictures in an all-in-English context while measuring event-related brain potentials (ERPs).

Participants were asked to press a button when the third picture of a triplet belonged to the same semantic category as the first two, and another button when it belonged to a different category. Unbeknownst to them, in half of the trials, the gender of the third picture name in Spanish had the same gender as that of the first two, and the opposite gender in the other half. We found no priming in behavioural results of either semantic relatedness or gender consistency. In contrast, ERPs revealed not only the expected semantic priming effect in both groups, but also a negative modulation by gender inconsistency in Spanish-English bilinguals, exclusively. These results provide evidence for spontaneous and unconscious access to grammatical gender in participants functioning in a context requiring no access to such information, thereby providing support for linguistic relativity effects in the grammatical domain.

1. Introduction

Whorf (1956), one of the two fathers of linguistic relativity, famously suggested that the language(s) one speaks shapes the way one thinks. The questions underpinning the linguistic relativity debate are questions such as: Does language shape thought? Is language encapsulated or does it interact with other cognitive processes? If so, what is the nature of these interactions and what properties of language bring these interactions to bear?

Scholars have misinterpreted Whorf's thesis as a formulation close to linguistic determinism, a far stronger claim that language may cause changes in basic physiological processes, e.g., that of visual perception. Pinker (1995), for instance, stated that "no matter how influential language may be, it would seem preposterous to a physiologist that it could reach down into the retina and rewire the ganglion cells" (p. 63). Over the past two decades, however, the linguistic relativity hypothesis has been resurrected in a milder form, perhaps closer to the original thinking of Sapir and Whorf (see, Gentner & Goldin-Meadow, 2003; Gumperz & Levinson, 1996). Additionally, interactive models developed by McClelland & Rumelhart (1981), and Humphrey, Riddoch, & Price (1997) and a recent, more direct application of these models to linguistic relativity by Lupyan (2012), offer working hypotheses as regards the cognitive mechanisms underpinning language-thought interactions. The label-feedback hypothesis, in particular, proposes that language is highly interconnected with

other cognitive processes such as vision and categorisation and influences other functional networks in a top-down fashion³.

Several recent studies have highlighted areas where lexical and grammatical information may affect cognitive processes other than language. Lexical characteristics of languages have been shown to affect colour perception in behavioural (Athanasopoulos, 2009; Franklin *et al.*, 2008; Özgen, 2004; Roberson *et al.*, 2005) and neurophysiological (Clifford *et al.*, 2010; Liu *et al.*, 2010b; Thierry *et al.*, 2009) investigations. A number of studies have also reported effects of language in the domain of spatial representation and event conceptualisation where speakers exhibit differences in event description and recollection (Bowermann & Choi, 1991; Majid *et al.*, 2004; Papafragou & Selimis, 2010) or even show different gaze patterns when exploring videos depicting events (Flecken, 2010). Studies investigating differences in grammatical number expression (e.g. languages with classifiers systems) have also suggested alteration of object classification (Athanasopoulos, 2007; Lucy, 1992b; Saalbach & Imai, 2007; Zhang & Schmitt, 1998).

One interesting feature of some languages, which offers an appropriate test case for linguistic relativity, is grammatical gender. This feature, present in many of the world's languages, forces all nouns to be assigned to, most commonly, two or three classes: *masculine* and *feminine* and/or *neuter* (Corbett, 1991).

Grammatical gender is of particular interest for two reasons: (a) when it is absent, it cannot be replaced by other lexicalisation patterns (unlike classifiers,

³ The label-feedback hypothesis is very recent and has only thoroughly been implemented in the context of label effects on object perception and categorization. However, there is no reason to believe that the feedback processes it outlines do not arise for grammatical features such as gender.

for instance), and (b) its assignment is arbitrary except in the case of natural gender (male/female distinction).

With regard to point (a), for instance, Chinese requires the use of a marker before every quantified noun as in ‘*yi zhang zhi piao*’ (a [FLAT OBJECT] bank note), and English can sometimes do the same as in ‘a piece of paper’ or ‘a flock of sheep’. By contrast, grammatical gender, when absent from a language, cannot be replaced by any combination of words.

Regarding point (b), the French word for ‘sentry’ (*une sentinelle*), for instance, is feminine, but rare must have been women sentries; a *toaster* is masculine in French (*un grille-pain*) but feminine in Spanish (*una tostadora*); in German a woman is feminine (*die Frau*) but a girl is neuter (*das Mädchen*). Even diachronically, the gender of nouns can change: the old word for girl in Polish used to be feminine (*ta dziewczyna*) but nowadays it is neuter (*to dziewczę*). In other words, both within a language and cross-linguistically, the relation between grammatical gender and word meaning appears to escape logic.

Studies investigating grammatical gender to date, have essentially focused on potential links between grammatical gender and object categorisation using (a) the voice-attribution paradigm (Bassetti, 2007; Forbes *et al.*, 2008; Sera *et al.*, 1994; 2002), (b) common noun–proper noun associations (Boroditsky *et al.*, 2003; Phillips & Boroditsky, 2003), (c) semantic ratings and adjective associations (Boroditsky *et al.*, 2003; Flaherty, 2001), or (d) a combination of the above methods (Ramos & Roberson, 2010; Vigliocco *et al.*, 2005).

Unfortunately, in all of these cases, the interpretation falls short of establishing the source of effects at an abstract level, disconnected from language itself. As Pinker (2007) puts it: “Speakers of different languages tilt in different directions in a woolly task, rather than having differently structured minds” (p. 148). The most recent and perhaps strongest evidence of grammatical gender-driven relativity comes from a study by Cubelli and colleagues (2011), who minimised the possibility that participants could use language as a strategy by using a non verbal semantic task on pictures. However, not all studies investigating implicit effects of grammatical gender on object categorisation have reported overwhelming evidence for such effects (see, Kousta *et al.*, 2008) and several have led to mixed results (Bassetti, 2007; Sera *et al.*, 2002).

The greatest limitation of studies conducted so far in this field is their reliance on behavioural measurements. Indeed, as vigorously argued by Pinker (2007), behavioural evidence is open to contamination by explicit and/or idiosyncratic strategies used by participants to resolve the tasks, a process that is likely to solicit language processing (e.g., inner speech, sub-vocal rehearsal of instructions, covert denomination of objects, lexical access, etc.). If language access is prompted by the task at hand, then nothing can be said of the spontaneity of this effect. What is needed then is a method, which detects spontaneous access to grammatical gender representations without explicit involvement of language and not merely inferred from behavioural observations (cf., Pinker, 2007).

In the present study, we asked participants to decide whether the third of a series of three objects presented one-by-one on the screen belonged to the same

semantic category as the two first ones. Semantic relatedness amongst the three objects was manipulated along with a covert manipulation of grammatical gender consistency (**Table 1**). We predicted that semantic incongruence would result in a modulation of the N400 wave, a negative-going potential with an average peak latency of 400 ms post-stimulus and known to reflect semantic integration mechanisms (Kutas & Hillyard, 1980; 1984). On the other hand, and critically, we hypothesised that grammatical gender inconsistency may modulate the Left-Anterior Negativity (LAN), an ERP marker of morphosyntactic processing (Friederici *et al.*, 1993; Friederici & Jacobsen, 1999; Hahne & Friederici, 1999; Thierry *et al.*, 2003). If such results were obtained, it would mean that grammatical gender is retrieved automatically and unconsciously rather than strategically and consciously during object categorisation.

Table 1. Example of experimental conditions.

Picture primes	Targets	Gender Congruency	Semantically Relatedness
tomato celery	asparagus	+	+
tomato celery	carrot	-	+
tomato celery	truck	+	-
tomato celery	bike	-	-

2. Methods & Materials

Participants. Participants were 16 Spanish native speakers with English as a second language (L2) (henceforth Spanish-English bilinguals). One participant was eliminated due to insufficient data quality. The 15 remaining participants (9 female, age: 32.6 ± 1.981 ; 6 male, age: 29.3 ± 2.028) were included in the final analyses. All the participants learned English at least in primary and secondary school in Spanish speaking countries and were living in the UK at the time of testing. Table 2 summarises participants' language experience and self-assessed proficiency in L1 and L2. At the time of testing, the participants were using L2 slightly more than L1, due to their immersion context but the difference in self-reported use was not significant. Proficiency was however significantly higher in L1 than L2 ($z = -2.49, p < .05$).

Table 2. Characteristics of Spanish-English participants

Measure	Mean	Standard Error
L1 self-rating (10 pt scale)	9.7	(0.1)
L2 self-rating (10 pt scale)	8.5	(0.4)
Daily L1 usage (%)	44.6	(6.5)
Daily L2 usage (%)	55.4	(6.5)
Age of L2 acquisition (years)	10.2	(1)
Length of Immersion (months)	52.6	(12.3)

Twenty native speakers of English who all reported that they were monolingual also took part in the experiment as a control group. Three of them were eliminated due to insufficient data quality. The 17 remaining participants (9 female, age: 19.57 ± 0.53 ; 8 male, age: 19.25 ± 0.16) were included in the final analyses. All the participants were right-handed and had normal or corrected-to-normal vision.

Materials. We selected 288 black-and-white line drawings from Snodgrass & Vanderwart (1980) and Székely and colleagues (2004). Pictures were grouped in 96 triads, such that the name of two first pictures had the same gender and they belonged to the same semantic category and the third picture was either from the same or a different semantic category and their name either had the same or the opposite gender as the two others. The 96 triads could therefore be split into 48 semantically related and 48 semantically unrelated associations, or 48 gender consistent and 48 gender inconsistent associations, providing 24 triads per individual experimental conditions.

Procedure. After participants filled out a questionnaire about their language learning background and self-assessed proficiency in L1 and L2, they were tested individually in a quiet room. They were seated in front of a computer monitor (CRT make 19", 100 cm from the screen) on which picture stimuli were displayed within a viewing angle of 8 degrees and handed a response box. The participants were instructed to press a given button if the three pictures of a triad belonged to the same semantic category and another button if not.

Participants were never told about the covert gender manipulation. On each trial, a fixation cross was presented for 1000 ms, followed immediately by the first prime for a duration of 600 ms, then the second picture appeared after a blank screen of 250 ms duration, for a duration 600 ms. Then the target (third picture) appeared after a variable interval randomly selected between 300 and 500 ms in steps of 50 ms, in order to cancel offset effects. The target remained on screen until participants responded. Five practice trials preceded the experimental trials. All experimental instructions were provided in English. The

order of blocks was counterbalanced across participants and the presentation of items was randomised within each block.

Electrophysiological Recording. The EEG was continuously recorded at a rate of 1kHz from 64 Ag/AgCl electrodes placed according to the extended 10–20 convention. Two additional electrodes were attached above and below the left eye and on either side of the left and right eye in order to monitor for eye-blinks and horizontal eye movements. Cz was the reference electrode during acquisition. Impedances were maintained below 5 k Ω for all 64 electrodes and below 10 k Ω for vertical electrooculogram electrodes. EEG signals were filtered off-line using a 30 Hz low pass zero phase shift digital filter.

Behavioural Data Analysis. Given that the task performed by participants was a semantic relatedness task, accuracy was not informative regarding access to gender information, especially if we consider that some semantic associations were far fetched due to the necessity of creating a fully counter-balanced experimental design. Indeed, some of the triads may have seemed unrelated to the participant when they were considered related by the experimenters and vice versa. Therefore, we considered reaction times irrespective of response accuracy and we did not consider potential differences in accuracy arising between groups or conditions. However, RTs shorter than 250 ms and differing by more than 2.5 standard deviations from the average RT in each condition and participant were individually discarded. A two-way ANOVA by participant was conducted on the RTs with semantic relatedness (related, unrelated), grammatical gender consistency (consistent, inconsistent) as within-subject

factors and group (Spanish–English bilinguals, English monolinguals) as between-subject factor.

Electrophysiological data analysis. Eye-blink artefacts were mathematically corrected using the algorithm provided in Scan 4.4™. The algorithm is derived from the method advocated by Gratton, *et al.*, (1983). Note that eye-blink occurred mostly after the response was made as a consequence of special instruction given to the participants. ERPs were then computed by averaging EEG epochs ranging from -100 to 1000 ms after stimuli onset. Baseline correction was applied in relation to 100 of pre-stimulus activity and individual averages were re-referenced to the global field power produced over the entire scalp. ERPs time-locked to the onset of target pictures were visually inspected and mean amplitudes were measured in temporal windows determined based on variations of the mean global field power measured across the scalp (Picton *et al.*, 2000). Four components were identified as expected. The P1 and N1 were maximal at parietal sites and were measured in the 100–150 ms range for the P1 and 170–230 ms for the N2, the N400 was maximal on central sites (Cz) and was measured in the 300–400 ms window. Finally, the Left Anterior Negativity, strongest at left anterior recording sites, was measured in the 380–600 ms window. Peak latencies were measured at sites of maximal amplitude (PO8 for the P1 and N1, CZ for the N4; FT9 for the LAN) and mean ERP amplitudes were measured in regions of interest around the sites of maximal amplitude (O1, PO3, PO7, O2, PO4, PO8 for the P1 and N1; C1, CZ, C2, CP1, CPZ, CP2, P1, PZ, P2 for the N4; FT9, FT7, FC5, F7, F5, AF7 for the LAN). Note that we did not conduct a full-scalp analysis because the modulation of the ERP

components were predicted to occur in the regions of interest and therefore statistical analyses of ERP mean amplitude were conducted in sets of electrodes determined *a priori* based on the LAN and N400 literature (cf. introduction). Mean amplitudes and peak latencies were subjected to a mixed repeated-measures ANOVA with semantic relatedness (related, unrelated), grammatical gender consistency (consistent, inconsistent) and electrode (6 or 9 levels) as within-subject factors and group (Spanish–English bilinguals, English monolinguals) as between-subject factor. In addition, paired sample t-tests were conducted between the gender consistent and gender inconsistent conditions millisecond-by-millisecond to determine the onset of differences between conditions (using a linear derivation of the 6 electrodes used in the mean amplitude analysis).

3. Results

Behavioural data. Accuracy in the semantic relatedness task was overall high (79%) and was not studied by group or condition for the reason stated in the method section. Regarding reaction times, we found a significant main effect of group ($F(1, 30) = 11.4, p < 0.002, \eta_p^2 = .28$) but we did not find any significant effect of semantic relatedness ($F(1, 30) = 0.824, p > 0.1$) or gender consistency ($F(1, 30) = 0.010, p > 0.1$) and no significant interactions between factors.

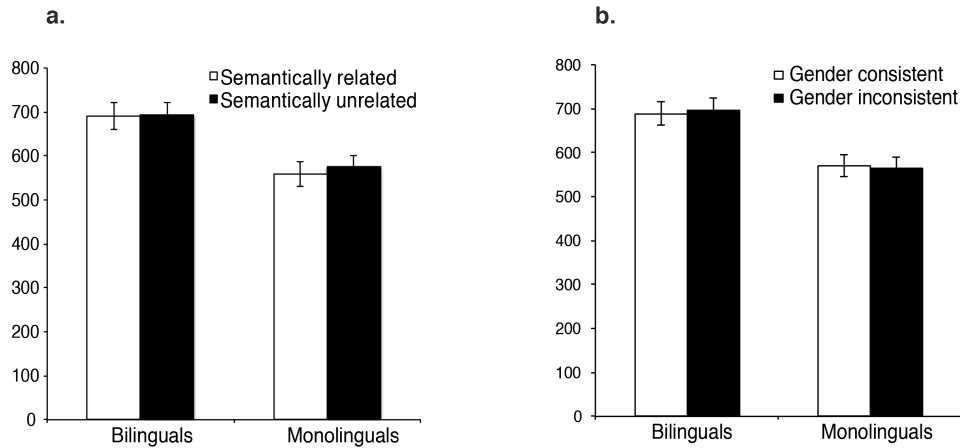


Figure 1. Plot of Reaction Times on correct trials only for the Spanish-English Bilinguals and English Monolinguals, showing no effect of condition. (A) Semantic conditions, (B) Gender conditions. Error-bars depict s.e.m

Electrophysiological data. N1 and P1 were unaffected by experimental conditions in either of the participant groups. As expected, the N4 was maximal over the centroparietal electrode sites and peaked at 361 ms on average (**fig. 2**). There was a significant main effect of semantic relatedness on N400 mean amplitude between 300 and 400 ms ($F(1, 30) = 40.74, p < .0001, \eta_p^2 = .576$) such that the N400 was less negative in the semantically related than unrelated conditions (Bonferroni, $p < 0.001$). There was no interaction between groups and the other experimental factors (**fig. 2 A & B**).

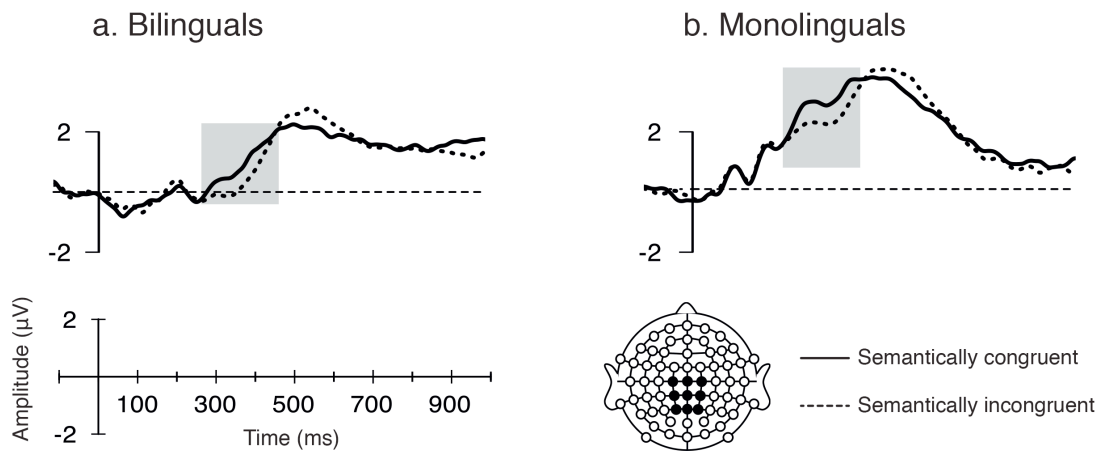


Figure 2. ERPs elicited in the semantic related and semantic unrelated conditions. (A) Spanish-English Bilinguals, (B) English Monolinguals. Linear derivation of electrodes C1, CZ, C2, CP1, CPZ, CP2, P1, PZ, P2.

The LAN was maximal over left frontal regions and peaked at 559 ms on average (**fig. 3**). There was a significant main effect of gender consistency ($F(1, 30) = 23.5, p < 0.0001, \eta_p^2 = .439$) such that LAN amplitudes were more negative in the gender inconsistent than gender consistent condition (Bonferroni, $p < 0.001$). This main effect was qualified by a significant interaction between gender and group ($F(1, 30) = 4.97, p < 0.05, \eta_p^2 = .142$). *Post-hoc* analyses revealed that the effect of gender was present for the Spanish–English bilinguals ($F(1,14) = 29.5, p < 0.001$, Figure 3 A) but not for the English monolinguals ($F(1,16) = 3.1, p > 0.05$, Figure 3 B).

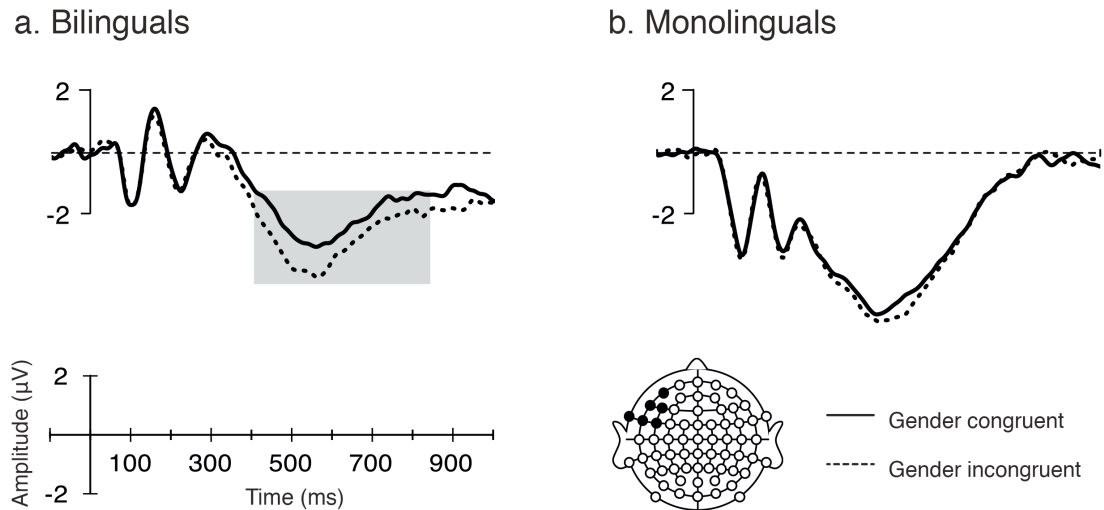


Figure 3. ERPs elicited in the gender consistent and gender inconsistent conditions. (A) Spanish-English Bilinguals, (B) English Monolinguals. Linear derivation of electrodes FT9, FT7, FC5, F7, F5, AF7.

Figure 4 plots the p -value (negative log of 10, for presentation purposes) of the t -tests carried out millisecond by millisecond on the difference between gender consistent and gender inconsistent conditions. The difference between conditions became significant 388 ms after stimulus onset in the Spanish-English bilinguals and remained so until the end of the analysed epoch (700 ms) and was never significant for more than 30 ms in the English monolinguals (Rugg, Doyle, & Wells, 1995).

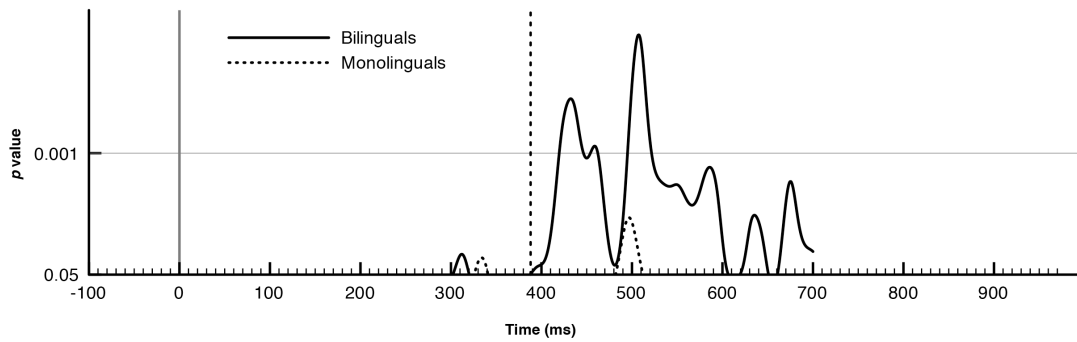


Figure 4. Plot of the p-value ms by ms for the LAN (negative log of 10). Dashed vertical line at 388 ms indicates when the effect becomes reliably significant.

4. Discussion

The aim of the present study was to determine whether some features of language affect other cognitive processes such as object categorisation.

Grammatical gender is a feature of some languages that has received considerable amounts of attention in linguistics and psycholinguistics but to our knowledge its effect on object categorisation has never been established based on measures of brain activity.

Unlike Cubelli *et al.* (2011) who used a similar experimental design, we found no behavioural effect of semantic relatedness of grammatical gender in Spanish–English bilingual participants. However, Cubelli *et al.* (2011) (a) used pairs rather than triads of pictures and (b) tested their participants in an Italian-speaking or Spanish-speaking environment (Italian and Spanish students tested at the university of Padova and Granada, respectively):

1. We used three pictures instead of two in order to (1) load participants' working memory, thereby increasing task difficulty and therefore reducing the likelihood of participants having enough executive resources to work

out the hidden manipulation and (2) maximise the explicit and implicit priming effects (i.e., experimental sensitivity) due to the more consistent baseline produced by the first two pictures;

2. The speaking environment has been shown to have a tangible impact on the language mode of individuals (Elston-Guttler, Gunter, & Kotz, 2005; Grosjean, 1998). Our participants were tested in an all-in-English context during and outside the experimental session. It must be noted that Welsh is also spoken in the region of North-Wales but that exposure to Welsh is rare to very rare if individuals are not actively seeking it, the medium of conversation being essentially English.

Nevertheless, we found the predicted semantic priming effect on N400 ERP amplitude (Kutas & Hillyard, 1980; 1984), showing that semantic priming amongst the picture triads was present even though it did not manifest itself behaviourally.

Critically, in addition to the semantic relatedness effect, we found a grammatical gender consistency effect in the ERP data exclusively in the Spanish-English bilinguals, manifesting itself as a LAN modulation and showing that these participants extracted gender information while engaged in a semantic categorisation task requiring no such information. Since participants were tested in an all-in-English context, were never made aware of the gender manipulation, never reported being aware of it after debriefing, and since gender consistency had no behavioural effect, we interpret this result as

evidence that access to gender information was implicit and unconscious (Kutas & Hillyard, 1980; 1984; Thierry & Wu, 2007; Wu & Thierry, 2010). This result indicates that grammatical gender is spontaneously retrieved during semantic processing of pictures even though lexical-semantic processing was not explicitly required (Strijkers *et al.*, 2011).

This result could be interpreted in terms of mere spreading of activation leading to the activation of grammatical gender representation even though accessing gender was irrelevant. Similarly, grammatical gender has been shown to affect picture naming cross-linguistically in bilingual word production (Lemhöfer, *et al.*, 2008) and there is good evidence that it is transferred in bilingualism (Ganushchak, Verdonschot, & Schiller, 2011). However, we note that most studies having brought to light such spontaneous effects of grammatical gender retrieval have used tasks that rely heavily on language activation. This was not the case in the current study since participants were not required to name pictures or retrieve any verbal information to perform semantic categorisation. In addition, the absence of a behavioural effect in our study suggests that spreading activation alone is not sufficient to account for the pattern of results obtained.

Irrespective of the fact that we did not find any behavioural effects, our conclusions are similar to those of Cubelli *et al.* (2011) with the added dimension that such effect is probably encountered on an unconscious level. Altogether our results suggest that object conceptual retrieval and categorisation are unconsciously affected by language-specific syntactic information, such as grammatical gender, even when such information is task-

irrelevant. Similar results of access to task-irrelevant semantic features were obtained by Yee, Ahmed and Thompson-Schill (2012). The demonstration of this phenomenon in the grammatical domain supports the view that language substantially interacts with other cognitive processes, and further highlights the critical role of language in shaping the way humans process reality and the world around them.

This conclusion is inconsistent with the modularity of language hypothesis (Chomsky, 2000; Fodor, 1975; 2008) and rather suggests that the organisation of information at the cortical level relies heavily on interconnectivity and interactions amongst distributed cell assemblies (Humphreys *et al.*, 1997; Martin, 2007; McClelland & Rumelhart, 1981; Pulvermüller, 1999). The data also lend support to the linguistic relativity hypothesis, and its newest development (Lupyan, 2012), by showing that semantic features of objects are spontaneously retrieved together with semantically irrelevant information such as syntactic gender and this information likely contributes to participants' mental representations of these objects.

3. Conclusion

While language does not necessarily determine thoughts, and while thinking may be possible without the aid of language, it nonetheless provides a ready basis of information for the purposes of classifying the world into meaningful categories (Lucy, 1997). To date, this observation has been empirically demonstrated primarily in the domain of colour (Clifford *et al.*, 2010; Liu *et al.*,

2010b; Regier & Kay, 2009; Thierry *et al.*, 2009). The current study shows that humans may automatically utilise grammatical categories such as gender when asked to make judgements about semantic relationships unrelated to the grammatical categories in question. The fact that we have found such effects in the domain of grammatical gender is particularly important, since previous empirical attempts to address the Whorfian question in this domain used methods and task instructions, mostly based on behavioural measures, that might promote strategic use of grammatical gender categories. Future studies will shed more light on the locus of this effect as well as patterns of brain connectivity, and establish whether it generalises to other language-specific properties, e.g., compound words, classifiers, and highly grammaticised language features, such as tense and aspect.

Chapter 5

Seeing objects through the language glass:
Unconscious effects of labels on perception

This paper is published as: Boutonnet, B., Dering, B., Viñas-Guasch, N., & Thierry, G. (in press). Seeing objects through the language glass. *Journal of Cognitive Neuroscience*.

Abstract

Recent streams of research support the Whorfian hypothesis according to which language affects one's perception of the world. However, studies of object categorisation in different languages have heavily relied on behavioural measures that are fuzzy and inconsistent. Here, we provide the first electrophysiological evidence for unconscious effects of language terminology on object perception. While English has two words for cup and mug, Spanish labels those two objects with the word '*taza*'. We tested native speakers of Spanish and English in an object detection task using a visual oddball paradigm, while measuring event-related brain potentials. The early deviant-related negativity elicited by deviant stimuli was greater in English than in Spanish participants. This effect, which relates to the existence of two labels in English versus one in Spanish, substantiates the neurophysiological evidence that language-specific terminology affects object categorisation.

1. Introduction

The question of language-thought interactions has recently become a major topic of interest in cognitive neuroscience. It has become essential because of the debate on language encapsulation and on the potential effects of language on other cognitive processes (Chomsky, 2000; Fodor, 1975; 2008). The linguistic relativity hypothesis has undergone several interpretations since its inception by Whorf (1956). One early (misleading) interpretation of the hypothesis contends that language *determines* thought and therefore that without language thought is impossible. In light of compelling evidence that high-level cognitive operations are possible without language, this position has simply become untenable (e.g., number cognition in primates (Gallistel, 1989), infants (Feigenson *et al.*, 2004) and in languages which do not have a complex lexicalised number system (Gordon, 2004)). On the other hand, recent theoretical reconceptualisations (Gentner & Goldin-Meadow, 2003; e.g., Gumperz & Levinson, 1996) have put forward non-deterministic versions of the hypothesis, according to which language *influences* (rather than determines) thought. The linguistic relativity debate has therefore moved towards the question of interaction between language representations and perception rather than that of determinism (Lucy, 1992a). However, this reconceptualization lacks psychological and physiological underpinning.

Here, we aimed at testing the validity of the most recent theoretical take on the Whorfian hypothesis, which does away with a “strong/weak” distinction (Klemfuss, *et al.*, 2012; Lupyan, 2012) and offers researchers clearer working hypotheses regarding language-thought interactions. For instance, based on

interactive-processing models such as those developed by McClelland and Rumelhart (1981), the *label-feedback* hypothesis (Lupyan, 2012) proposes that language is highly interconnected with other cognitive processes such as vision and categorisation and that it produces transient modulations of on-going perceptual (and higher-level) processing. Whorfian effects can therefore arise from interactions amongst distributed brain regions, as in the case of prefrontal areas preparing the visual cortex to perceive particular dimensions of stimuli before they are actually displayed (Lamme & Roelfsema, 2000). This model therefore allows for non trivial linguistic relativity effects to arise but is not tied in a deterministic view where perceptual areas are functionally *structured* by language (for an exhaustive explanation of the hypothesis and a review of the experimental literature see Lupyan (2012)).

Previous studies have highlighted areas where lexical and grammatical information affect domain-general cognitive processes. For instance, lexicalisation constraints on spatial representation and event conceptualisation (e.g., focus on manner vs. end-point of motion) have been shown to affect speakers' event description and recollection (Bowermann & Choi, 1991; Majid *et al.*, 2004; Papafragou & Selimis, 2010) or to elicit different gaze patterns when exploring scenes depicting events (Flecken, 2010). Studies investigating grammatical number (i.e., language with classifier systems) reveal a tendency to categorise objects on the basis of substance rather than shape when classifiers put the focus put on substance (Lucy, 1992b; Saalbach & Imai, 2007; Zhang & Schmitt, 1998). In a similar vein, grammatical gender has also been shown to affect speakers' object

categorisation in covert gender assignment tasks (Bassetti, 2007; Forbes *et al.*, 2008; Kurinski & Sera, 2010; Sera *et al.*, 2002), judgement and adjective-association tasks (Boroditsky *et al.*, 2003; Phillips & Boroditsky, 2003), and priming paradigms (Boutonnet *et al.*, 2012; Cubelli *et al.*, 2011). Finally, differences in colour terminology have been shown to affect colour perception in behavioural (Athanasopoulos, 2009; Franklin *et al.*, 2008; Özgen, 2004; Roberson *et al.*, 2005) and neurophysiological (Athanasopoulos, Dering, Wiggett, Kuipers, & Thierry, 2010b; Clifford *et al.*, 2010; Liu *et al.*, 2010b; Thierry *et al.*, 2009) investigations.

Despite the evidence in favour of the existence of Whorfian effects, it remains that studies in the field have mostly relied on behavioural measures. The problem in that such measures are open to contamination by explicit strategies used by participants to resolve the tasks, a process likely to involve language processing. Here, following neurophysiological investigations in the domain of colour (Franklin *et al.*, 2008; Gilbert *et al.*, 2006; Liu *et al.*, 2010b; Roberson, Pak, & Hanley, 2008; Thierry *et al.*, 2009), we investigated whether language-specific terminology also constrains object categorisation (Gilbert *et al.*, 2008). For instance, Thierry and colleagues (2009) recorded ERP correlates of colour change detection in Greek-English bilinguals who have two colour terms for blue (*ble* - 'dark blue' and *ghalazio* - 'light blue') and a control group of English monolinguals. They found that native speakers of Greek exhibited a greater visual mismatch negativity (vMMN) elicited by blue deviants than English controls. Based on this paradigm we chose to extend the evidence from the domain of colour perception to that of object

categorisation. We chose drinking vessels because they have been examined thoroughly in previous cross-linguistic naming studies (Ameel, Malt, Storms, & Van Assche, 2009; Ameel, Storms, Malt, & Sloman, 2005; Pavlenko & Malt, 2010). These studies suggest that bilingual speakers' categorical boundaries shift through exposure to their second language –a phenomenon that has also been reported for colours (Athanasopoulos, *et al.*, 2010b).

We recorded brain potentials from Spanish and English native speakers while they performed an object detection task within an oddball paradigm to test the extent to which unconscious aspects of visual object processing are modulated by one's language.

Spanish differs from English in the way some objects are labelled. While English has two words to refer to a cup and a mug, Spanish only uses one label for these two objects: *taza*. In this experiment, participants were presented with three stimuli within an oddball paradigm (one of high local probability, i.e., standard, and two of low local probability, i.e., deviants). Participants were instructed to detect a particular deviant stimulus, or target (a bowl) in each of two experimental blocks. In one block, the non-target deviant was a cup and the standard was a mug and in the other block, the non-target deviant was a mug and the standard was a cup.

We expected non-target deviants to spontaneously elicit a deviant-related negativity (DRN) regardless of a response from the participants (Csibra, Czigler, & Ambro, 1994; Czigler, Balázs, & Pató, 2004; Czigler, Balázs, & Winkler, 2002; Turatto, Angrilli, Mazza, Umiltà, & Driver, 2002; Winkler,

Czigler, Sussman, Horváth, & Balázs, 2005). Because of the terminological difference between English and Spanish, we expected that the change from cup to mug would elicit a greater DRN in English than Spanish participants.

2. Materials & Methods

Participants. Participants were 13 native speakers of Spanish (10 female, 3 male; $M_{AGE} = 21 \pm 1.6$) tested in Spain and 14 native speakers of English tested in Wales (8 female, 6 male; $M_{AGE} = 20 \pm 0.6$). Spanish participants were recruited from a database filtered to have a level no higher than A2 in English and a daily use of English lower than 5%. As part of the normal education curriculum in Spain, all Spanish participants received some exposure to English but all reported having a limited knowledge of the language as well as a rare use of it. None of the Spanish participants had spent more than two weeks in an English-speaking country. The language usage background data used to filter the database were collected from self-reports from the participants prior to entry in the database. Some of the Spanish speakers were also fluent in Catalan. This was not considered a problem since Catalan and Spanish are matched with respect to object denomination for cups and mugs. Some of the English participants reported having basic knowledge of other languages (including Spanish) but had self-reported very low proficiency and were not using any of their other languages on an everyday basis.

Materials. Three grey-scale photographs of a cup, a mug, and a bowl subtending approximately 8° of visual angle were presented in the middle of a white background square in the centre of a CRT monitor.

Procedure. Participants viewed two blocks of 450 stimuli. Within each block, a standard stimulus was presented with a high local probability (either a cup or a mug, 80%). Deviant stimuli, presented with a low local probability, were either to be ignored (a mug or a cup, depending on the nature of the standard, 15%) or to be reported (bowl target, 5%). Presentation order was pseudo-randomised such that two deviants or targets never appeared in immediate succession, and there were at least three standards in a row between two deviants. Stimuli were presented for 300 ms with a random variable inter-stimulus interval of 400, 450, 500, 550, 600 ms, averaging to 500 ms. Participants were instructed to detect the target object (bowl) by pressing a button on a response box as quickly as possible. Block order was fully counterbalanced between participants.

Electrophysiological recording. Electrophysiological data was recorded in two different labs. The Spanish participants were tested in Barcelona, Spain (Pompeu Fabra University). EEG was recorded (BrainVision Recorder 1.10™) in reference to the left mastoid electrode at the rate of 1 kHz from 34 tin electrodes placed according to the 10–20 convention. Impedances were kept below 5 kΩ for electrodes on the cap and below 10 kΩ for external electrodes. The English participants were tested (NeuroScan 4.4™) in Bangor, Wales (Bangor University). EEG was recorded in reference to left mastoid electrode at the rate of 1 kHz from 34 Ag/Cl electrodes placed according to the 10–20 convention. All impedances were kept below 5 kΩ for electrodes on the cap and below 10

k Ω for external electrodes. Both datasets were analysed using BrainVision Analyzer 2™. EEG activity was filtered offline with a high-pass 0.1 Hz filter (slope 12 dB/oct) and a low-pass 30 Hz filter (slope 48 dB/oct).

Data analysis. Accuracy scores and reaction times were submitted to independent samples t-tests between groups (t_1 and t_2 , respectively). Eye-blinks were mathematically corrected using the Gratton, Coles & Donchin (1983) algorithm provided in Brain Vision Analyzer 2™, and epochs with activity exceeding $\pm 75 \mu\text{V}$ at any electrode site were automatically discarded. Epochs ranged from -100 to 600 ms after stimulus onset. Baseline correction was performed in reference to pre-stimulus activity and individual averages were re-referenced to the left and right mastoid offline. ERPs time-locked to the onset of the pictures were visually inspected and mean amplitudes were measured in temporal windows determined based on variation of the mean global field power measured across the scalp (Picton *et al.*, 2000). ERPs elicited by standard stimuli were averaged across blocks as were ERPs elicited by deviants, therefore comparisons between standard and deviants did not reflect inherent perceptual differences between cups and mugs but only the deviancy effect.

Potential perceptual differences between the cup and mug objects were also investigated by analysing amplitude and latency of the P1 peak from ERPs computed from standard stimuli separately for each of the two experimental blocks. The P1 was maximal at parietal sites and was measured in the 100–150 ms range. Mean amplitude and latency of the P1 collected from a linear derivation of the 5 electrodes of interest (PO1, PO2, O1, OZ and O2) were submitted to a 2 within- x 2 between-subject ANOVA with standard object

(cup/mug) as a within-subject factor and language group (Spanish/English) as a between-subject factor.

The DRN was defined as the earliest modulation of the negative component following the P1 over occipital recording sites. DRN analysis was conducted on individual ERPs elicited by standards and non-target deviants and was maximal over the parietoccipital scalp and studied in the 145–180 ms range at electrodes PO1, PO2, O1, OZ, and O2, predicted to be the electrodes of maximal sensitivity for the effect measured (Liu *et al.*, 2010b; Thierry *et al.*, 2009). Mean amplitudes of ERPs from standard and deviant stimuli were subjected to a mixed repeated measures ANOVA with deviancy (deviant/standard) and electrode (5 levels) as a within-subject factors and language group (Spanish/English) as a between subject factor. In addition, paired sample t-tests were conducted between the standard and deviant conditions millisecond-by-millisecond to determine the onset of differences between conditions (using a linear derivation of the 5 electrodes used in the mean amplitude analysis).

Furthermore, the latency of the N1 elicited by non-target deviants was compared to that of the N1 elicited by the standards, measured at the electrode of maximal amplitude (O2). Peak latencies were submitted to a 2 within- x 2 between-subject ANOVA with deviancy (standard/deviant) as a within-subject factor and language group (Spanish/English) as a between-subject factor.

Since some native speakers of Spanish were also Spanish-Catalan bilinguals, we investigated potential differences in attention allocation between groups by comparing ERPs elicited by mug standards and bowl targets on the one hand

and cup standards and bowl targets on the other hand, because these comparisons always involved objects that have different names in both of the languages. P1s and DRNs elicited by *cup*, *mug* and *bowl* (in identical time windows and the same electrodes as the analyses above) were subjected to repeated-measures ANOVAs with object (cup-bowl/mug-bowl) as within-subject factor and language group (Spanish/English) as a between-subject factor. Because of the very high-level of repetition involved in the oddball paradigm use here, we expected potential differences in attention to have a negligible impact on basic object discrimination as indexed by DRN. We therefore expected to find no interaction between object type and group in these comparisons.

3. Results

Behavioural data. Accuracy in the bowl detection task was above 90% in all participants and blocks, ($M_{\text{ENGLISH}} = .94 \pm .02$; $M_{\text{SPANISH}} = .93 \pm .02$). There was no significant differences between groups on target detection accuracy nor RTs ($t_1(25) = .62, p > .05$; $t_2(25) = .29, p > .05$).

Electrophysiological data.

Critical comparison: Standard (cup/mug) versus passive deviant (cup/mug)

As expected non-target deviants elicited a greater DRN as compared to standards. This difference was qualified by a significant main effect of deviancy ($F(1, 25) = 10.3, p < .05, \eta_p^2 = .29$) with deviant stimuli eliciting more negative amplitudes than standard stimuli in the DRN window. The effect of deviancy

further interacted with language group ($F(1, 25) = 4.9, p < .05, \eta_p^2 = .16$) such that the deviancy effect was of significantly greater magnitude in English than Spanish participants

(Figs. 1A and 1B).

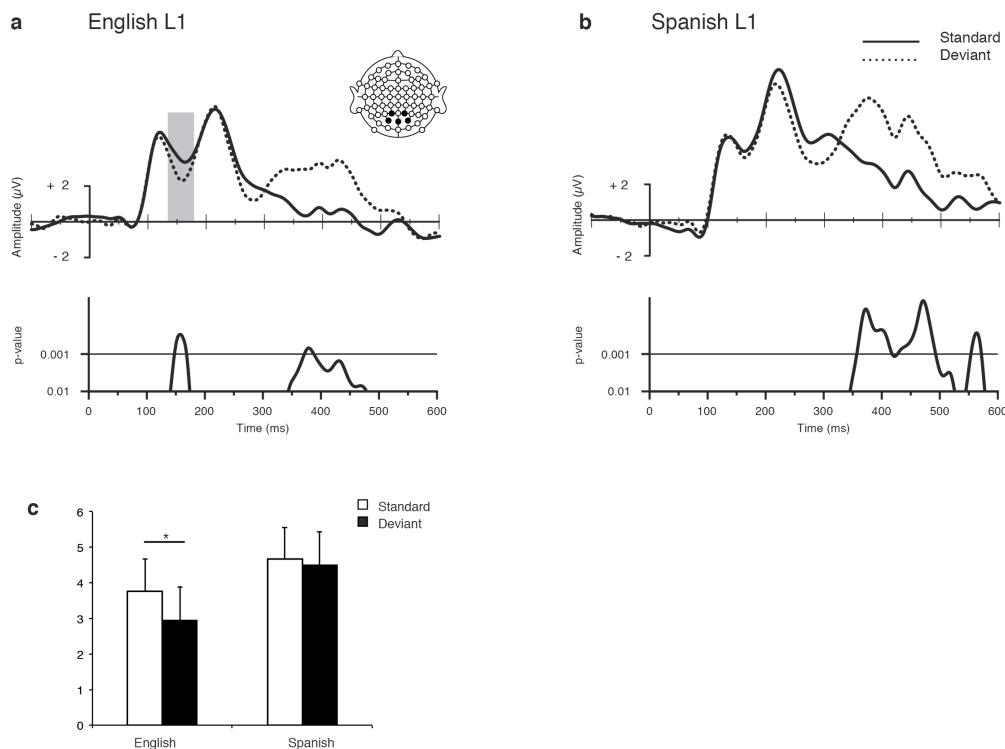


Figure 1. Event-related brain potentials elicited by standard and deviant stimuli averaged across blocks. ERPs and plots of p-value of differences between conditions in (A) Native speakers of English and (B) Native speakers of Spanish. (C) Plot of DRN mean amplitude. Waveforms correspond to linear derivation of electrodes PO1, PO2, O1, OZ, O2. Error bars depict s.e.m.

Post hoc test showed that there was no significant DRN effect in the Spanish group ($F(1, 12) = .46, p > .05, \eta_p^2 = .04$) but a significant effect in the English group ($F(1, 13) = 16.31, p = .001, \eta_p^2 = .56$, Fig. 1C). Furthermore, there was no

significant difference between standard and deviant conditions at any point in time in the DRN window in the Spanish participants, but standard and deviant conditions differed significantly from 135 – 177 ms in the English group (lower part of **Figs. 1A and 1B**). In order to reduce the risk of Type I errors and given the high levels of autocorrelation of ERP time series, we followed the method advocated by Guthrie and Buchwald (1991) where only sequences with a minimum of 12 consecutive significant t-test were considered (see, for instance, Kuipers and Thierry (2011)). Latency analyses of the DRN revealed no significant differences between group or condition in the window of interest ($F(1, 24) = 1.53, p > .05, \eta_p^2 = .06$). ERPs elicited by standard stimuli in each of the two blocks considered separately (**Fig. 2**) displayed significant differences in P1 mean amplitude (F_1) and latency (F_2) between cup and mug ($F_1(1, 24) = 5.76, p < .05, \eta_p^2 = .19; F_2(1, 24) = 17.56, p < .001, \eta_p^2 = .42$). Critically, these effects did not interact with participant group ($F_1(1, 24) = 1.29, p > .05, \eta_p^2 = .05; F_2(1, 24) = 3.2, p > .05, \eta_p^2 = .12$).

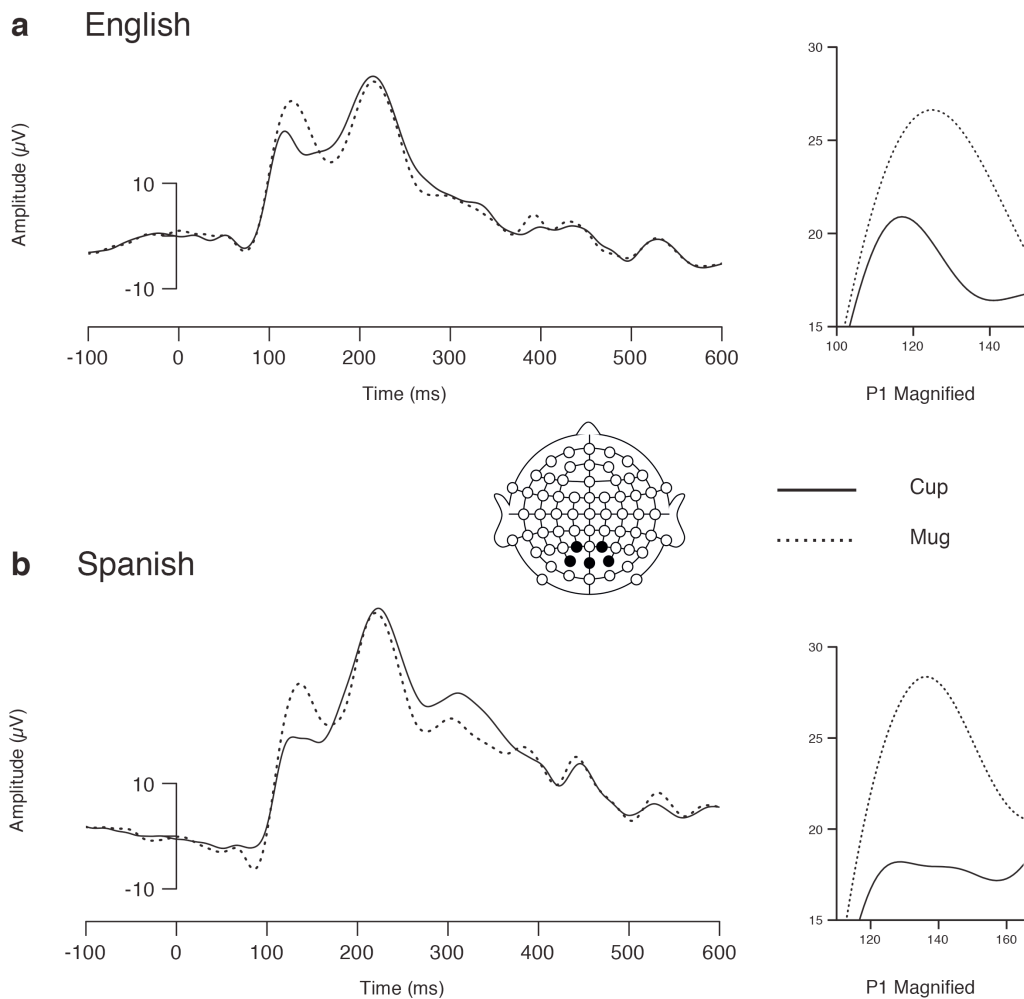


Figure 2. Event-related brain potentials elicited by cup and mug standards in each of the two experimental blocks. (A) Native speakers of English and (B) native speakers of Spanish. Waveforms correspond to linear derivation of electrodes PO1, PO2, O1, OZ, O2.

Control comparison: Standard (cup/mug) versus target (bowl)

ANOVAs on the P1 (**Fig. 3**) revealed a significant effect of object type in both the mug vs. bowl comparison ($F(1, 25) = 50.32, p < .0001, \eta_p^2 = .69$) and the cup vs. bowl comparison ($F(1, 25) = 40.28, p < .0001, \eta_p^2 = .62$). Critically, there was no interaction between language group and object type in either comparisons (both $ps > .1$).

ANOVAs on the DRN (**Fig. 3**) revealed significant effect of object type in both the mug vs. bowl comparison ($F(1, 25) = 40.28, p < .0001, \eta_p^2 = .62$) and the cup vs. bowl comparison ($F(1, 25) = 48.57, p < .0001, \eta_p^2 = .66$). Again, there was no interaction between language group and object type in either comparisons (both p s $> .1$).

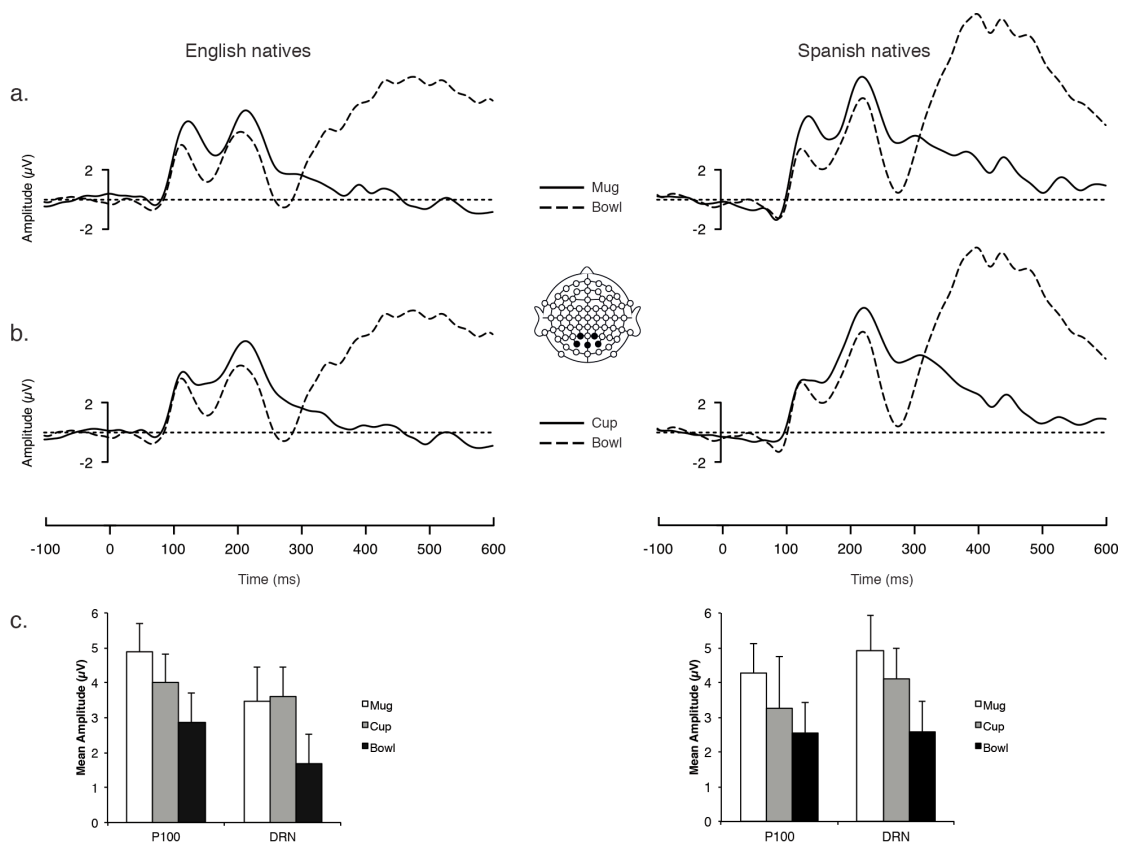


Figure 3. Event-related brain potentials elicited by (A) mug standards and bowl targets and (B) cup standards and bowl targets. (C) Plot of P1 and DRN mean amplitudes in both participant groups. Waveforms correspond to linear derivation of electrode PO1, PO2, O1, OZ, O2. Error bars depict s.e.m.

4. Discussion & Conclusions

This study tested potential effects of language-specific terminology on early stages of visual perception and categorisation based on the analysis of spontaneous modulations of the P1/N1 event-related brain potential complex. In a design controlling for perceptual features of the objects presented, ERPs successfully distinguished standards and deviants within the N1 range in native speakers of English but not speakers of Spanish who name both these objects using the same noun. Moreover, when comparing the P1 elicited by the two objects presented as standards in each of the blocks, ERP differences were indistinguishable between groups.

The N1 range of ERPs is thought to index stages of visual processing beyond categorical discrimination (Dering, Martin, Moro, Pegna, & Thierry, 2011; Thierry, Martin, Downing, & Pegna, 2007a). Indeed, categorical effects have been reported in the domain of face processing in the P1 range and even earlier (Seeck, Michel, Blanke, Thut, Landis, & Schomer, 2001; Seeck *et al.*, 1997; Thierry, Martin, Downing, & Pegna, 2007a; 2007b). Therefore, since it occurs beyond the P1 range, the DRN effect found here concerns relatively sophisticated levels of visual object processing –probably relating to object identity resolution. Critically, however, the DRN occurred before the temporal window in which lexical representations are considered to be accessed. Indeed during practiced picture naming Costa, Strijkers, Martin, and Thierry (2009) and Strijkers, Holcomb, and Costa (2010) have established that lexical access occurs between 180 and 200 ms after picture onset. Here significant differences were observed as early as 145 ms after picture onset. In addition, as shown by

Strijkers *et al.* (2011) lexical access appears to be substantially delayed until ~350 ms after stimulus onset when there is no requirement to name the pictures (see also, Blackford, Holcomb, Grainger, & Kuperberg, 2012). This was indeed the case here since participants were asked to press a button when they saw a specific object and not instructed to name them. Thus, the influence of language-specific terminology on object processing does not merely result from online interaction with processes underlying lexical access. In other words, our finding is not simply an effect of language on language.

We report the N1 modulation recorded here as a DRN rather than a vMMN (the visual counterpart of the auditory MMN (Czigler *et al.*, 2002; Winkler *et al.*, 2005)) because the vMMN *proper* is supposedly only elicited by visual stimuli presented outside the focus of attention, e.g., in peripheral vision rather than fixation (Clifford *et al.*, 2010). However, (a) the latency of the DRN effect we reported here is similar to that previously reported in vMMN studies (Pazo-Alvarez *et al.*, 2003); (b) like our effect, the vMMN has a parietoccipital topography with a right hemispheric predominance. Since the DRN in the present study (peak time ~160 ms at electrode O2) peaked substantially earlier and was observed at a different scalp location than N2 modulations elicited by overt cognitive control (Folstein & Van Petten, 2007), we interpret this effect as an index of automatic, pre-attentional and, crucially, pre-lexical cognitive mechanism (Costa *et al.*, 2009; Strijkers *et al.*, 2010; Strijkers *et al.*, 2011).

The P1 results further suggest that Spanish and English participants perceptually discriminated cup and mug pictures in a similar fashion. These two objects are indeed ostensibly different and P1 amplitude has been shown to

distinguish different object types previously (e.g., Dering *et al.*, 2011; , Thierry, *et al.*, 2007a). Therefore the DRN effect observed in the N1 window cannot be explained by differences arising at more elementary stages of perceptual analysis preceding the N1 window. Furthermore, we consider the absence of between-groups differences in the P1 range to be of fundamental importance since, they could underpinned by differences in cultural background or ethnic origin or even genetic factors and would therefore invalidate our results as merely stemming from different perceptual grooming in different environments.

Differences between groups in the P1 range could have been expected since our group has already reported such differences in a previous study of colour perception (Thierry *et al.*, 2009). However, it must be noted that the relationship between colour terminology and P1 measurement was not trivial in that it did not yield a P1 amplitude by language group interaction. Expecting a reduction or cancellation of P1 differences between cups and mugs in the Spanish participants here would assume that perceptual differences between a cup and mug are even more subtle than perceptual differences between two neighbouring shades of blue, which have been shown to occur between 100 and 200 ms after stimulus onset (Fonteneau and Davidoff, 2007). We contend that cups and mugs are more discriminable at a perceptual level (at least by shape, size, and luminance) than are two discs of the same size and colour saturation, differing exclusively by their relative luminance. For example, people will argue indefinitely about colour names at the green-blue or the navy-indigo border but the same individuals will hardly argue as to what differentiates a mug and a cup shape. Therefore, it is reasonable to assume that P1 differences indexing early

perceptual distinctions should effectively discriminate cups and mugs in both groups but that orientation responses measured by the DRN would be selectively affected by language terminology.

The fact that differences occur only in the N1 range and based on standard/deviant comparisons is essential to demonstrate an effect of language terminology on high-level perceptual processing. Additionally, these differences arising beyond the P1 range are consistent with an interactional account of linguistic relativity effects (Lupyan, 2012) since basic perception need not be changed for such effects to arise.

Our experimental design also allowed us to investigate potential attentional differences between the Spanish-Catalan speakers and English monolinguals. Indeed, one could argue that the interaction on the DRN could be a result of better inhibition/monitoring mechanisms in the bilinguals. As suggested by our results, this was not the case since, when the items both had a different label in Spanish and English, the DRN elicited between target and standard had the same magnitude in the two groups. If Spanish participants had different attentional skills, and if such skills were generically reflected in DRN modulation, we would have expected the interaction observed in the critical comparison (mug / cup) to carry over to the case of comparisons with the target (bowl).

To our knowledge, this is the first neurophysiological demonstration of a relationship between native language and spontaneous object identity discrimination during visual perception, which goes beyond the observation of

overt effects on object categorisation (Ameel *et al.*, 2005; 2009; Pavlenko & Malt, 2010). Furthermore, these findings generalise the linguistic relativity effects previously reported in the case of colour perception (Franklin *et al.*, 2008; Liu *et al.*, 2010b; Thierry *et al.*, 2009) to the domain of object identity processing (Gilbert *et al.*, 2008) (arguably affecting higher-level cognitive representations). Overall, our results are incompatible with the view that language is functionally encapsulated in the human brain and fundamentally independent of e.g., visual cognition (Chomsky, 2000; Fodor, 1975; 2008; Pinker, 1995). On the contrary, they support an interactive conceptualisation of the brain where language is highly integrated and can modulate ongoing cognitive processes such as object categorisation and perception (Lupyan, 2012). Future studies will determine whether the effects reported here are confined to interactions within the left hemisphere (Franklin *et al.*, 2008; Mo, Xu, Kay, & Tan, 2011; Regier & Kay, 2009; Roberson *et al.*, 2008) and the extent to which they are adaptable over time (Athanasopoulos *et al.*, 2010b).

Chapter 6

Compound words prompt arbitrary
semantic associations in
conceptual memory

This paper is submitted as: Boutonnet, B., McClain, R., & Thierry, G. Compound words prompt arbitrary semantic associations in conceptual memory.

Abstract

Linguistic relativity theory has received empirical support in domains such as colour perception and object categorisation. It is unknown however, whether relations between words idiosyncratic to language impact representations and conceptualisations. For instance, would one consider the concepts of horse and sea as related were it not for the existence of the compound seahorse? Here, we investigated such arbitrary conceptual relationships using a non-linguistic picture relatedness task in participants undergoing event-related brain potential recordings. Picture pairs arbitrarily related because of a compound and presented in the compound order elicited N400 amplitudes similar to unrelated pairs. Surprisingly, however, pictures presented in the reverse order (as in the sequence horse – sea) reduced N400 amplitude significantly, demonstrating the existence of a link in memory between these two concepts otherwise unrelated. These results break new ground in the domain of linguistic relativity by revealing predicted semantic associations driven by lexical relations intrinsic to language.

1. Introduction

The Whorfian hypothesis that language may influence other cognitive processes has recently become a major topic in psycholinguistics and neuroscience, probably because evidence in this area is directly informative as regards long-standing debates on language encapsulation (Chomsky, 2000; Fodor, 1975; 2008).

Over the past two decades, the Whorfian hypothesis has undergone a significant revival. First, and in agreement with early criticism of the hypothesis, a deterministic view of linguistic relativity has been dismissed. If anything, in light of compelling evidence that high-level cognitive operations are indeed possible without language, this position becomes untenable (Feigenson *et al.*, 2004; Gallistel, 1989; Gordon, 2004). Subsequent developments of the hypothesis have proposed a non-deterministic reading according to which language influences thought without necessarily determining it. The most recent theoretical development of the Whorfian hypothesis, by Lupyan (2012), offers clear working hypotheses regarding the type of processes which may lead to Whorfian effects. The *label-feedback* hypothesis (Lupyan, 2012) proposes that language is highly interconnected with other cognitive processes such as categorisation and that it produces transient modulations of on-going neural processing at different functional levels.

Recent studies have highlighted areas where lexical and grammatical information affect domain-general cognitive processes. For example, colour terminology has been shown to influence categorical perception of colour in monolingual and bilingual speakers (Athanasopoulos *et al.*, 2010b; Franklin *et*

al., 2008; Gilbert *et al.*, 2006; Liu *et al.*, 2010b; Roberson *et al.*, 2008; Thierry *et al.*, 2009). Language-specific lexicalisation of events and spatial representation has been shown to affect speakers' perception, recollection and even gaze patterns when exploring pictures and videos depicting events (Bowermann & Choi, 1991; Flecken, 2010; Majid *et al.*, 2004; Papafragou & Selimis, 2010). Finally, several studies provide evidence that grammatical number and gender expression can alter speakers' object perception and categorisation (Athanasopoulos & Kasai, 2008; Boroditsky *et al.*, 2003; Boutonnet *et al.*, 2012; Cubelli *et al.*, 2011; Saalbach & Imai, 2007).

Despite accumulating evidence in favour of linguistic relativity, critics remain unconvinced. One of the greatest limitations of most previous studies on the hypothesis is their heavy reliance on behavioural measures. Behavioural evidence leaves open the possibility of a contamination by explicit top-down strategies soliciting language processing during tasks that are misconstrued as nonverbal, which, in the context of an investigation of unconscious and automatic effects of language on other cognitive processes, is insufficient, and could boil down to a mere effect of language on language ("thinking for speaking" (Slobin, 2003), see also Pinker, 2007).

While most of the literature on linguistic relativity has focused on effects at the interface between language and other cognitive processes, here, for the first time, we study idiosyncratic relations existing within language and test whether arbitrary relations between lexical entities induce arbitrary association in conceptual space. We tested whether the existence of compound words lead to associations between the concepts referred to by their morphemes (e.g.,

between *sea* and *horse* in *seahorse*). Several studies have provided evidence that the compound's morphemes are decomposed and accessed individually in reading (Andrews, Miller, & Rayner, 2004; Duñabeitia, Laka, Perea, & Carreiras, 2008; Fiorentino & Poeppel, 2007; Hyönä, Bertram, & Pollatsek, 2004; Koester, Gunter, & Wagner, 2007). In addition, several neuroimaging studies have identified combinatorial processes involved in comprehension of simple noun-noun (Graves, Binder, Desai, Conant, & Seidenberg, 2010) and metaphorical phrases (Forgács, Bohrn, Baudewig, Hofmann, Pléh, & Jacobs, 2012). Finally, Gagné and Spalding (2009) have suggested that the integration of compounds rely on the same combinatorial processes involving both psycholinguistic and conceptual knowledge.

We hypothesised that language-specific combinations afforded by compound words lead to the establishment of associations in conceptual-semantic memory. Using event-related brain potentials (ERPs) and a non verbal picture relatedness judgment task, we tested whether the relation between the two morpheme of a compound would lead to a reduction of N400 amplitudes (Kutas & Hillyard, 1980; 1984) elicited by the second picture of a pair. We expected pairs of pictures related conceptually or via the existence of a compound to reduce N400 amplitude as compared to random pairs of the same pictures created whilst avoiding conceptual relationships.

2. Methods & Participants

Participants. Participants were 16 (9 female, 7 male, age: 21.9 ± 0.9) native speakers of English and students of the School of Psychology at Bangor University. They were offered course credits for their participation in the study that was approved by the ethics committee of Bangor University.

Materials. We selected 51 compound words (e.g., sandcastle) and a prototypical picture for each morpheme embedded within them (e.g., a picture for sand, and one for castle). Altogether, 102 highly recognisable photographs were selected from online image databases. The pictures were arranged into 4 fully rotated experimental conditions: semantically related (Related), related via a compound and in the compound order (Compound), related via a compound but in the reverse order (Reversed), and semantically unrelated (Unrelated). Picture stimuli subtending approximately 10° of visual angle were presented on a white background in the centre of a 19" CRT monitor.

Experimental Condition	Picture Names		Association Strength
Related	SEA	FISH	.036 (.009)
Compound	SEA	HORSE	.045 (.008)
Reversed	HORSE	SEA	.008 (.003)
Unrelated	SEA	CAKE	.003 (.002)

Table 1. Experimental conditions, example of names of pictures in each of the conditions, and Edinburgh Associative Thesaurus (Kiss, Armstrong, Milroy, & Piper, 1973) mean relatedness scores. S.e.m is given in brackets.

Association strength of all picture name pairs was extracted from the Edinburgh Associative Thesaurus (Kiss *et al.*, 1973) and a repeated-measures Analysis of Variance (ANOVA) was conducted to test for differences in

semantic relatedness between picture names across conditions (see **Table 1**). There was a significant main effect of condition ($F(3, 147) = 9.6, p < .0001, \eta_p^2 = .17$). Post hoc comparisons showed that the Related and Compound conditions were significantly more related than the Reversed and Unrelated conditions (Bonferroni, $ps < .05$), respectively, but did not differ significantly from each other. Critically, the Reversed pairs were not significantly more related than the Unrelated pairs (Bonferroni, $p > .1$).

Procedure. Participants signed a consent form to take part in the study that was approved by the Ethics Committee of Bangor University. They were tested individually in a quiet room and instructed to press a given button when two consecutive pictures were related and another button when they were unrelated. Participants were not informed about the presence of pairs derived from compound words and were instructed to focus on evaluating the conceptual relatedness of the pictures. On each trial, a fixation cross was presented for 250 ms, followed immediately by the first picture for a duration of 500 ms, then the second picture appeared after a random variable inter-stimulus interval of 400, 450, 500, 550, 600 ms, averaging to 500 ms, and remained on screen for a duration of 3000 ms maximum or disappeared upon participant response. A blank screen with a duration of 500 ms on average separated each trial, in order to cancel offset effects. Block order was fully counterbalanced across participants and stimulus presentation was fully randomised.

Electrophysiological recording. The EEG was continuously recorded at a rate of 1kHz from 64 Ag/AgCl electrodes placed according to the extended 10–20 convention. Two additional electrodes were attached above and below the left

eye and on either side of the left and right eye in order to monitor for eye-blinks and horizontal eye movements. Cz was the reference electrode during acquisition. Impedances were maintained below 5 k Ω for all 64 electrodes and below 10 k Ω for vertical electrooculogram electrodes. EEG activity was filtered offline with a high-pass 0.5 Hz filter (slope 12 dB/oct) and a low-pass 30 Hz filter (slope 48 dB/oct).

Behavioural data analysis. Two separate repeated measures ANOVAs were carried out on RTs (F_1) and on accuracy (F_2) with the 4 conditions (related, compound, reversed, and unrelated) as within-subject factors.

Electrophysiological data analysis. Eye-blink artefacts were mathematically corrected using the algorithm provided in Scan 4.4™. The algorithm is derived from the method advocated by Gratton *et al.* (1983). Note that eye-blink occurred mostly after the response was made as a consequence of special instruction given to the participants. ERPs were then computed by averaging EEG epochs ranging from -100 to 1000 ms after stimuli onset. Baseline correction was applied in relation to 100 ms of pre-stimulus activity and individual averages were re-referenced to the global field power produced over the entire scalp. ERPs time-locked to the onset of target pictures were visually inspected and mean amplitudes were measured in temporal windows determined based on variations of the mean global field power (Picton *et al.*, 2000). Three components were identified as expected. The P1 and N1 were maximal at parietal sites and were measured in the 100–150 ms range for the P1 and 170–230 ms for the N1. The N400 was maximal over central sites and was measured in the 350–480 ms window. Peak latencies were measured at sites of

maximal amplitude (PO8 for the P1 and N1, Cz for the N4) and mean ERP amplitudes were measured in regions of interest around the sites of maximal amplitude (O1, PO3, PO7, O2, PO4, PO8 for the P1 and N1; F3, Fz, F4, FC1, FCz, FC2, C1, Cz, C2 for the N4). Note that we did not conduct a full-scalp analysis because the modulation of the ERP components were predicted to occur in the regions of interest and therefore statistical analyses of ERP mean amplitude were conducted in sets of electrode determined a priori. Finally, mean amplitudes subjected to a repeated-measures ANOVA with condition (Compound, Reversed, Related and Unrelated), anteriority (anterior, central, posterior) and laterality (left, centre, right).

3. Results

Behavioural data. Statistical analyses carried out on RTs revealed no significant differences between experimental conditions ($F_1(3, 42) = .44, p > .05, \eta_p^2 = .03$; **Fig 1**). Accuracy analysis revealed a main effect of condition ($F_2(3, 42) = 21.34, p < .001, \eta_p^2 = .60$). *Post-hoc* analyses found no significant differences between the Compound and Reversed conditions (Bonferroni, $p > .05$), but the Related and Unrelated conditions were significantly different from all other conditions, with Related leading to significantly lower accuracy scores than all other conditions ($M = .53 \pm .03$, Bonferroni, $p < .05$) and Unrelated leading to significantly higher accuracy scores than all other conditions ($M = .92 \pm .02$, Bonferroni, $p < .05$), as illustrated in **Fig 1**. Finally, accuracy significantly differed from chance (.5) in all conditions ($ps < 0.0001$) but the Related condition.

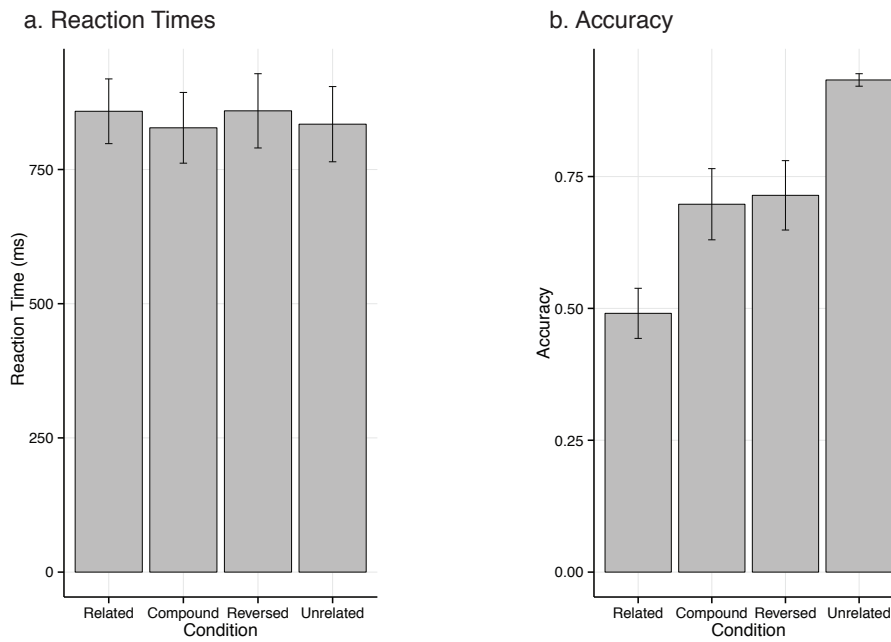


Figure 1. Plot of mean reaction times (RTs) and accuracy in the four experimental conditions. Error bars depict s.e.m.

Electrophysiological data. Statistical analyses carried out on N400 mean amplitudes revealed a main effect of condition ($F(3, 45) = 6.84, p < .001, \eta_p^2 = .31$), a main effect of anteriority ($F(2, 30) = 23.6, p < .001, \eta_p^2 = .61$), and a main effect of laterality ($F(2, 30) = 8.91, p < .001, \eta_p^2 = .37$). There was no significant interaction. *Post hoc* paired t-test revealed significant differences between Related and Unrelated ($t(15) = -3.04, p < .05$; **Fig. 2a**), between Related and Compound ($t(15) = -3.6, p < .05$; **Fig. 2b**) and between Reversed and Unrelated ($t(15) = 2.3, p < .05$; **Fig. 2c**).

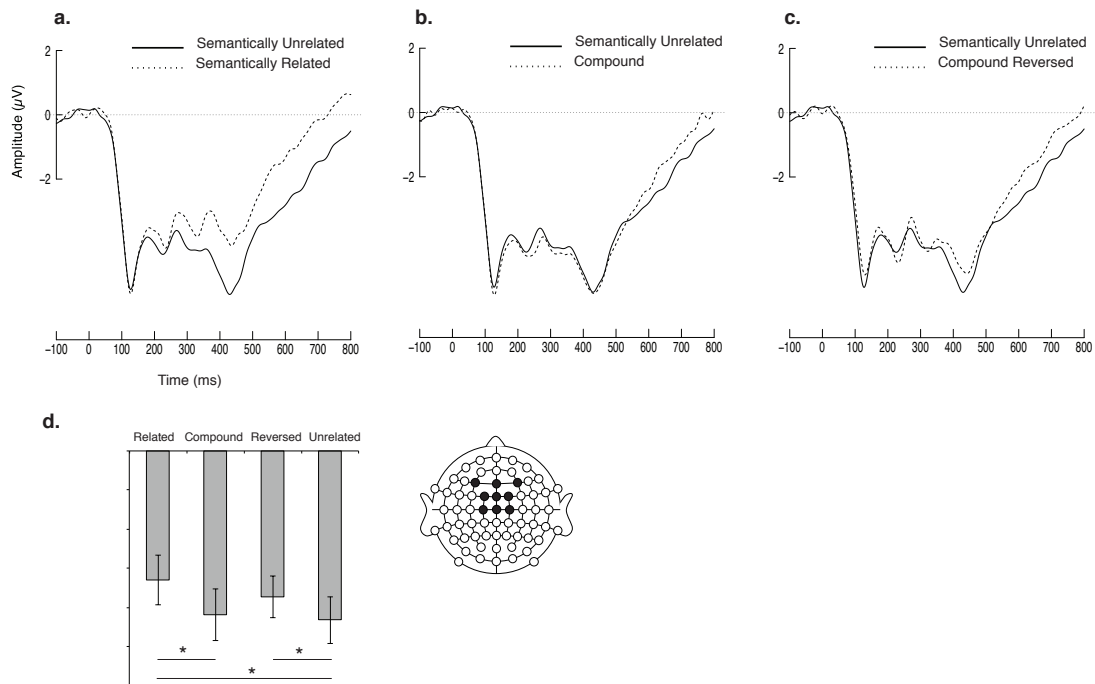


Figure 2. Event-related brain potentials elicited by the four experimental pairs averaged across blocks. (A) ERPs elicited in the Related and Unrelated conditions. (B) ERPs elicited in Related and Compound conditions. (C) ERPs elicited in the Reversed and Unrelated conditions. (D) Bar graph of mean N400 amplitudes in all experimental conditions. Waveforms correspond to linear derivations of electrodes F3, Fz, F4, FC1, FCz, FC2, C1, Cz, C2. Error bars depict s.e.m.

4. Discussion

This study investigated whether a phenomenon such as word compounding, which leads to artificially boosted lexical relations idiosyncratic to language, has consequences regarding the organisation of conceptual-semantic knowledge. Whereas ERPs elicited by pictures related because of the existence of a compound in the lexicon failed to reduce N400 ERP amplitudes, the same pictures presented in the reverse order significantly reduced N400 amplitudes as compared to semantically unrelated ones.

Although there was no difference in RTs between conditions, there was a significant difference in accuracy such that conceptually related pictures led to error rates not different from chance. Although judging the relatedness of pictures is not a difficult task *per se*, this result is not surprising given that the related pictures were not prototypically related (**Table 1**). This was due to the fact that we were not at liberty to select any related picture pair to serve as related pairs because this would have engendered uncontrolled differences in low-level visual differences, visual complexity (Dering *et al.*, 2011; Thierry *et al.*, 2007a), familiarity, prototypicality, etc. For this reason, all trials (correct or incorrect according to our predictions) went into the behavioural and ERP analyses in order to retain comparable trial sizes. However, because there were clear divergences between our and participants' expectations, we chose to explore the potential contribution of participants' judgements on N400 amplitudes. First, we calculated the proportion of related responses for each trial type. This relatedness ratio implemented into a linear regression model was found to be a significant predictor of N400 amplitude (N400 amplitude by Relatedness Ratio, $\Delta R^2 = .32$, $p < .01$) in the related condition only (results of the model are summarised in **Table 2 and fig. 3**).

Table 2. N400 amplitude by Relatedness Ratio regression output.

	ΔR^2	<i>B</i>	<i>SE B</i>	β	<i>P</i>
	0.32				
Constant		-9.7	2.53		<0.001
Relatedness Ratio		12.48	4.88	0.56	<0.01

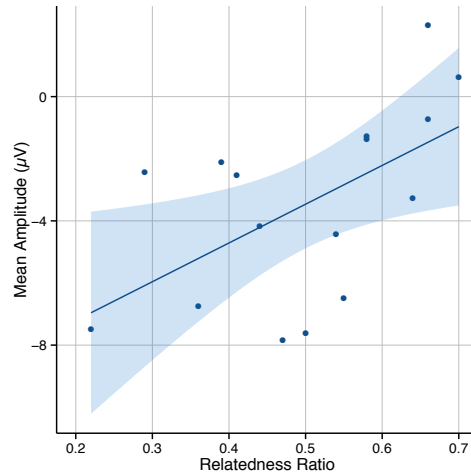


Figure 3. Scatter plot illustrating the relationship between N400 amplitude and Relatedness ratio. Line depicts linear regression shaded area depicts confidence interval.

As illustrated in figure 3, the result suggests that the more participants responded related on in the Related condition the less negative the N400 amplitude. Participants' subjective judgements did not influence any of the other experimental conditions, which means that the ERPs in the critical compound and reverse compound conditions are meaningful independently of participants' judgements. Additionally, this result also suggests that if more prototypical pairings had been available for the Related condition, N400 amplitudes would have been less negative and differences observed in the present data would have only been stronger. In any case, we can be reassured that semantic priming did take place, since we obtained highly significant N400 differences between the Related and Unrelated conditions, replicating classical N400 effects (e.g., Kutas & Federmeier, 2011). It is also important to note that the discrepancy between our judgements and the participants' judgements was only an issue for the Related condition since participants had accuracy scores

higher than 70% in all other conditions (cf., **fig. 1**). This analysis confirms one of the benefits of collecting electrophysiological data in more constrained experimental designs where behavioural measures lack sensitivity. Moreover, accuracy was intermediate between related and unrelated conditions in the Compound and Reversed conditions, which, in the absence of RT differences, could be interpreted as a sign that Compound and Reversed pairs were less related than related pairs but still more related than unrelated pairs.

ERPs, however, shed a very different light onto the underlying mechanisms involved. Although Compound and Reversed conditions were indistinguishable on the basis of behavioural data, N400 amplitude was only significantly reduced as compared to unrelated pairs for the latter condition. This result suggests that the lexical link between a compound's morphemes has consequences in terms of conceptual-semantic memory associations, since pairs of objects *a priori* unrelated like *horse – sea* seem to be easier to integrate than a completely unrelated pair (e.g., *horse – shell*) or the compound-ordered picture pair (e.g., *sea – horse*). This effect cannot be driven by the particular stimuli used in every condition, since the association scores for the Unrelated and Reversed conditions did not differ, and, as critically explained above, the specific pictures used in all the conditions were identical. The most conservative interpretation of this effect is thus that a link in conceptual space has been imposed by idiosyncratic language connections.

Surprisingly, however, the Compound condition, which had association ratings comparable to the related condition, and is, intuitively, the condition in which the lexical link would have been most obvious, failed to reduce N400

amplitude contrary to our hypothesis. We attribute this fact to interference arising from a conflict between the meaning of the compounds' morphemes considered separately and that of the compound word considered as a whole, reflected by an increase in N400 amplitude.

An alternative and less exciting account could be that as participants process compound-ordered pairs of pictures, they access in their lexicon the actual compound word whose lexical frequency is necessarily lower than that of the words depicted by the pictures considered individually. This would also be expected to result in greater negativity in the N400 range (Kutas & Federmeier, 2011). Indeed participants were presented with pictures and, despite the fact that they were not required to name them, it cannot be ruled out that lexical representations were accessed (albeit delayed as compared to lexical access during picture naming (Strijkers *et al.*, 2011)).

Irrespective of the explanation for a lack of semantic priming in the compound-ordered condition, the effect observed in the reversed-compound condition constitutes strong evidence that lexical relations imposed by language within the lexicon have implications for semantic memory organisation. Indeed, the significant reduction of N400 amplitudes when participants processed the second image of a compound pair presented in reversed order, demonstrates the existence of a conceptual link that does not exist for similarly unrelated concepts, which are not artificially related via compounding. Future studies will investigate how such effects differ between languages, since a numerous compounds are language-specific.

Chapter 7

General discussion

1. Summary of the main results

In this thesis I investigated the linguistic relativity hypothesis using ERPs and made the following three observations:

- Grammatical gender information is retrieved in a non-verbal semantic categorisation task using pictorial stimuli, although such information is irrelevant, unnecessary and, in fact, counter-productive, for successful performance in the task. This is interpreted as evidence that daily exposure to a language with grammatical gender, a feature that must be attended to for the production and understanding of nouns, becomes intimately linked with conceptual representation of objects.
- Lexical distinctions, brought about by language, influence the way objects are represented at the interface between perception and conceptualisation. Similar to previous research on colour distinctions, the presence of labels seems to skew visual stimulus categorisation.
- Links established in the lexicon by constructs idiosyncratic to language such as compound words affect the organisation of conceptual-semantic knowledge. Indeed, language exposure and the use of compound words, which “physically” link two otherwise unrelated words and concepts, lead to the emergence of conceptual relatedness.

Overall, our results provide strong support in favour of the hypothesis that language can affect other (arguably language independent) cognitive processes. At one end of the spectrum, such effect can manifest in a strong fashion, e.g., by modulating early visual processing (cf., Chapter 5). At the other end, Whorfian effects can manifest in a weaker guise, through activations of irrelevant (grammatical) features which have become intimately linked with conceptual representations due their repeated activation (cf., Chapter 4). The work presented in this thesis not only offers an investigation into the linguistic relativity hypothesis on different cognitive levels, but also confirms and extends previous findings in two of the most studied domains: grammatical gender and colour perception. In addition, the neurophysiological data provided sheds light onto the timing and cognitive make up of Whorfian effects. The discussion below focuses on an integrated account of the findings obtained in the three experimental chapters without rehearsing the individual points raised in their dedicated discussion.

2. Is the distinction between verbal and non-verbal processing meaningful?

The data presented in this thesis argue in favour of a highly dynamic and integrated conception of Whorfian effects, whereby language is fully integrated in the human brain and interacts closely with “non-verbal” processes.

Furthermore, depending on the processes involved, the influence of language can be relevant in a different way and to a different extent. This conclusion

questions the classical dissociation made between “verbal” and “non-verbal” processing. The repercussions of this point are twofold:

(1) In line with recent theoretical implementations of the linguistic relativity hypothesis in a wider psychological model, such as that proposed by the label-feedback hypothesis (Lupyan, 2012), verbal and non-verbal processes can be considered parallel and combined rather than separate. Indeed, I have argued earlier that it is precisely this distinction which has led to an explanatory inadequacy since considering those two processes as completely different commits most explanations to the pitfalls of determinism. Indeed, for effects of verbal on non-verbal processes to take place one must *modify* the other. Since Whorfian effects do not take such a form, they are often accounted for in terms of the weakest understanding on the LRH, which contends that effects of language on thought are limited to the production of language, thereby “ignoring” the strongest (and most interesting) effects.

(2) The evidence that language is highly integrated with other processes calls for a reconceptualisation, focussing on differential requirements brought about by task demands. If, for instance, a cognitive judgement relies heavily on visual properties of a presented stimulus, effects of terminology are more likely to be measured than if the task requires semantic judgements. Beyond this, general cognitive requirements (attention, memory and executive functioning) will modulate the overt or covert reliance on language strategies, thereby changing the manifestation of Whorfian effects.

3. Language is highly integrated in a not-so-hierarchical brain

The results presented in this thesis argue for a full integration of language with other “non-verbal” systems. Verbal and non-verbal processing are intimately and intrinsically integrated such that one influences the other and *vice versa* at different levels and interfaces, in keeping with fully integrated conceptions of the human brain (Elman, 2004; Kiefer & Pulvermüller, 2012; Mahon & Caramazza, 2008; McClelland & Rumelhart, 1981; Pulvermüller, 1996; 1999; 2012). Indeed, the results reported here strongly argue against a modular conception of language (Chomsky, 2000; Fodor, 1975; 2008). Early proposals of the integration of language, such as captured by Elman (2004) (based on Rumelhart’s (1979) conception of words and their meaning) suggest that words themselves do not have meaning but rather that they are cues to meaning such that words are considered as stimuli triggering mental states (activate representations which are dependent on distributed brain regions) and that it is those mental states that are meaningful.

“My approach suggests that comprehension, like perception, should be likened to Hebb’s (1949) palaeontologist, who uses his beliefs and knowledge about dinosaurs in conjunction with the clues provided by the bone fragments available to construct a full-fledged model of the original. In this case, the words spoken and the actions taken by the speaker are likened to the clues of the palaeontologist, and the dinosaur, to the meaning conveyed through these clues.” (Rumelhart, 1979).

The best example of such a view follows from the recent developments of the hypothesis of embedded cognition, which contends that the brain is a wide interconnected network of cell assemblies without individual distinctions of functional specialisation. When brain cells are led to communicate in a functionally relevant way, patterns of activity influence each other whether they are involved in aspects of language processing, or perception, or else (see, Borghi & Pecher, 2011 for a recent review; Mahon & Caramazza, 2008; Pulvermüller, 1996; 1999). Pulvermüller (2012), for instance, suggests that language is “woven into action” at the level of the brain, whereby hearing or reading action words (such as verbs) activate areas of the motor cortex that are related to these specific actions. These data show that different, and sometimes quite long-distance, neural circuits interact and influence each other during the processing of word and sentence meaning.

It is important to note, however, that the data does not directly address the claims made by embodied accounts of cognition nor does it attempt to argue in favour of it. Indeed, the results are just as compatible with accounts for embodied cognition as they are against, since the core finding is the prevalence of brain interconnectivity during the processing of meaning (Mahon & Caramazza, 2008; Borghi & Pecher, 2011, Barsalou, 2008 for recent reviews}. Indeed, although proponents of disembodied accounts of cognition reject the fact that meaning is embodied in the brain’s modal systems for perception, action and introspection, they claim that cognition is grounded by *interaction*, whereby concepts are abstract and symbolic but may still interact with sensory

and motor systems so that they may instantiate online perceptual processing (Mahon & Caramazza, 2008; 2009).

Overall, criticisms such as Pinker's (Pinker, 1995; 2007) that behavioural evidence is weak or squarely ineffective in establishing linguistic relativity effects, are challenged by the electrophysiological data presented in this thesis, since ERP effects are not directly under the control of the participant and weakly affected, if at all, by strategy, especially when these effects are registered in response to implicitly manipulated linguistic properties. Additionally, as explained in Chapter 2, there is a growing number of studies providing evidence that the brain is not exclusively a feed-forward, hierarchical system, but rather structurally organised in a way that interactions between areas dedicated to "early" and "late" processing can occur at all times (Churchland *et al.*, 1994; Freeman, 2000; Lamme & Roelfsema, 2000; Mesulam, 1998).

4. On different levels of modulations

The results reported in this thesis present three different kinds of effects, which tap into different levels of cognition. A DRN (Chapter 5), strongly related to visual processing, a LAN (Chapter 4), strongly associated with the processing of morpho-syntactic features, and an N400 modulation (Chapter 6), usually indexing semantic integration. This spread of indices could be seen as difficult to interpret regarding the nature and cognitive locus of Whorfian effects. One could even wonder whether any of these effects qualifies as indices of Whorfian effects or whether they could be considered by-products of the cognitive

operations encouraged by the tasks involved and therefore have nothing to say about effects of language on other cognitive processes. I argue that the variety of manifestations of Whorfian effects should not undermine their interpretation as such since, as proposed by the label-feedback hypothesis (Lupyan, 2012) and consistent with an integrated view of human cognition, such variety of guises is in fact expected. If language is to interact dynamically and on-line with other cognitive processes, its effects and modulations should be taking place on any of the levels where language characteristics have some relevance. This point highlights the importance of neurophysiological data because of their high sensitivity and relative independence *vis-à-vis* behavioural measures.

Furthermore, RT and accuracy measures not only lack online temporal information (shedding light onto intermediary stages of processing), but these measures might be too variable or insensitive to detect subtle modulations at particular levels of processing.

5. On late LAN interpretation

In chapter 4, I report that when Spanish speakers performed a categorisation task on triplets of related or unrelated pictures, a LAN was elicited when the grammatical gender of the name of the last picture in a series of three mismatched the gender of the two other picture names. I have reported the effect as a LAN, since it was elicited in relation to a morphosyntactic property and took the form of a negativity over the left-frontal region of the scalp. The activation of gender features was taken as evidence that grammatical gender influences on-line categorisation of objects. However, in this section I wish to

offer a more speculative interpretation of this component based on the implications of the interactional account of human cognition that these data support. Traditionally, the LAN is elicited in the context of sentence processing and upon presentation (visual or auditory) of a morphosyntactic violation (e.g., *she **go** to the swimming pool vs. she **goes** to the swimming pool) and tends to peak (i.e., reach its greatest negative amplitude) between the 300-500 ms time window, with a left-anterior distribution. The LAN elicited in my paradigm was significantly delayed. It reached significance from 388 ms and peaked at around 600 ms (~100 ms delay). If we consider the categorisation task that the participants performed, even if grammatical gender information was activated, it is unclear why this would result in a LAN modulation, since it is almost exclusively limited to linguistic processing and to contexts where explicit morpho-syntactic violations are presented. In other words, instead of a LAN one could conceptualise the effect found here as a late detection of the grammatical gender inconsistency which is entirely co-lateral to semantic evaluation and not generated by the same LAN generator normally involved during syntactic parsing. If we compare the timings of the N400 and LAN, we can see that the LAN started splitting after the N400 reached its maximal peak. I suggest that the LAN may be functionally related to the N400, such that it would index retrieval and assessment of morpho-syntactic features attached to object representations in a similar way as the N400 indexes semantic feature retrieval and assessment. Such an interpretation is in keeping with the idea behind the label-feedback hypothesis which contends that effects do not need to take the form of a physical wrapping but rather that Whorfian effects arise

functionally through on-line modulations. In fact, in a word-pair priming paradigm, Thierry, *et al.* (2003) found a LAN elicited by the second noun of a pair when it mismatched the grammatical gender of the first. However, their LAN significantly differed from classical LANs in terms of timing and topography from which they concluded that, “under these particular conditions ... grammatical gender incongruence seem to elicit ... modulations [that are similar to semantic incongruent conditions]” (Thierry *et al.* 2003, p.544). I believe that this observation calls for closer and dedicated examination in future experiments. It is also particularly interesting in the perspective of an interpretation of our results in terms of spreading activation (see below).

6. On Spreading Activation

An anonymous reviewer of Boutonnet *et al.* (2012) suggested that the LAN result could be interpreted in terms of mere spreading activation as such effects have previously been reported (e.g., Lemhöfer *et al.*, 2008). In the experimental chapter, I discussed the fact that in the study by Lemhofer *et al.* (2008) the task directly involved language processing and therefore it is expected that spreading of activation may trickle down to other language properties of the stimuli.

While a spreading activation account cannot be ruled out with regard to the present data, the task we used involved conceptual categorisation and, even though it is not incompatible with language co-activation, it arguably required no language process to be active. In fact, considering the available evidence on the time-course of lexical access, it is unlikely that lexical labels were retrieved, since lexical access is estimated to take place at around 200 ms during practice

picture naming but is significantly delayed until ~350ms when there is no intention to do so (Strijkers *et al.*, 2011). Additionally, and in light of the interpretation of the LAN that I have mentioned above, the effects found in this study are consistent with a functional link between grammatical gender information and other semantic representations of objects, where the morpho-syntactic “processor” is called upon to perform a feature-check in terms of grammatical gender because this property is connected to the semantic representation of the object. Although the gender-related negativity obtained over the left anterior regions shows sensitivity to a violation of gender context in the predicted direction, which is seemingly compatible with repetition priming, indiscriminate spreading of activation may have been expected to result in behavioural effects as well. The absence of the latter suggests that mechanisms of suppression are also at work and therefore spreading of activation alone cannot account for the pattern of results observed.

7. On the absence of (Whorfian) behavioural effects

It is often the case that ERP studies are criticised when significant modulations measured in the ERPs are not measurable in terms of –or incongruent with– behavioural effects. The argument is that the absence of behavioural results should cast doubts on the reliability of the ERP data. This critique is at least partly valid when studies investigate modulations of late components known to correlate strongly with reaction times (e.g., P300 family) or error rates (e.g., error-related negativity). However, unduly generalised, this line of reasoning makes the assumption that ERP data are a mere high tech addition to

behavioural measurements and do not provide independent evidence for the phenomenon under study. Naturally, I consider this position entirely flawed. ERPs index processes that may or may not influence RT and accuracy measures, since numerous brain processes do not result in measurable behavioural modulations and since processes which take the same overall time to unfold from stimulus presentation to response may have very different intermediary stages. Additionally, not all experimental designs lend themselves to the elicitation of behavioural effects, and this is particularly the case in the experiments I conducted here given the fact that they measure effects, which are (1) subtle and (2) linguistic in origin but occur at a non-linguistic level in tasks that do not involve linguistic stimuli. In the study presenting activation of grammatical gender features, I have explained that the absence of behavioural effects may be due to suppression mechanisms or that the gender context, which we chose to establish by presenting two picture primes, may have loaded participants' memory. As noted in the discussion of the experiment, Chapter 4, behavioural effects on a similar priming paradigm (using only two pictures) have been reported (Cubelli *et al.*, 2011), which suggest that behavioural effects can be elicited by covert manipulation of grammatical gender. However, the predicted ERP effects were obtained in terms of both semantic priming (N400 modulation) and grammatical gender priming (LAN modulation). In other experimental contexts, such as that of the oddball paradigm, no behavioural data is registered for any of the stimuli presented that require no response (only targets require a response, and even in this case responses can be omitted). In this case, the measure of choice is a modulation of N1 amplitude considered in

and of itself as the index of automatic change detection. There is no theoretical reason why behavioural data would be superficial in an oddball paradigm but required in the case of priming studies. However, it is expected that such an effect may also manifest behaviourally provided an appropriate experimental paradigm is used. For instance, based on previous research (Lupyan & Spivey, 2008), if participants perform a visual search with cups and mugs, it would be expected that the label benefit in the English participants would lead to faster detection rates compared to Spanish speakers. This question is even more important when we consider the experiment on compound words presented in Chapter 6, where one would expect significant behavioural effects since participants are actively engaged in the task. As discussed in the chapter the absence of behavioural effects is likely to result from the relatively poor levels of semantic relatedness in the “related” pairs, which were highly constrained because of the requirement to fully rotate the design, that is presented the exact same stimuli on all experimental conditions (thus controlling for properties such as visual complexity or lexical frequency). The ERP differences between conditions, however, speak for themselves and replicate N400 modulations observed in classical studies of semantic relatedness effects (see Kutas & Federmeier, 2011 for review). All together, in the context of the present studies and considering how Whorfian effects may arise in the brain, the absence of behavioural results only demonstrate the limitations of a purely behavioural approach. Considering Whorfian effects as an online modulation of on-going processes (Lupyan, 2012), it is expected that such effects may not necessarily have detectable behavioural consequences but that the online “thought-content”

would nonetheless be modified by language exposure. Such language exposure, could have, for instance, (1) arbitrarily established relationships between objects, (2) made object characteristics appear more or less salient, or (3) tune or prepare perceptual systems.

8. Online language effects on visual processing or effects of language on memory and decision processes?

In a recent experiment, Klemfuss, Prinzmetal, & Ivry (2012) issue a cautionary note on results provided by Lupyan & Spivey (2008; 2010b) suggesting that language modulates early visual processing. In a visual search paradigm, using distorted 2s and 5s, which could be seen as abstract symbols, Lupyan and Spivey (2008) found that when participants were instructed to consider the figures as rotated 2s and 5s, they were faster to detect the target. They conclude that ascribing meaning to abstract symbols leads to the activation of labels, which “prepare” visual processing areas and make them more efficient in such a task. Klemfuss *et al.* (2012) argue that although a dynamic account of online effects on visual processes arising from interactions between labels and visual areas is possible and appealing, the results could just as well be explained in terms of effects of language on “higher-level” processes such as memory and decision making. They claim that the facilitation effects measured when rotated 2s and 5s are considered numerals (essentially treated as meaningful via a label) could be due to a familiarity effect, which would in turn reduce processing load in terms of working memory. They claim that familiar shapes would be easier to retain in working memory, while unfamiliar shapes (when no label is associated with them) lack verbal representation in long-term memory and require their

representations to be encoded for each search trial. In an attempt to tap into potential memory effects, Klemfuss *et al.* (2012) designed a new task with a reduced load on working memory so as to make the two conditions (label / no label) more comparable. Instead of a fixation cross, the target was present in the centre of the screen so that storage in working memory would not be required to perform the task. They hypothesise that if language modulates visual levels it would still create a benefit in the cue condition. Their results suggest that ascribing meaning makes no difference in the cue condition, from which they conclude that language must not be affecting visual processing itself.

However, it is important to note that the cue condition used by these authors potentially changes the mode of operation by turning the cue and no-cue condition into two different tasks. Indeed, the no-cue condition can be considered a classic visual search paradigm, in which labels improve participants' performance, whereas the cue condition requires participants to operate on a pattern-matching basis, where the cue is serially compared to distractors, thus requiring less high-level processing. Under such conditions it is not surprising that influence from labels may be bypassed. It is a possibility that the authors themselves admit in passing but quickly dismiss (Klemfuss *et al.*, 2012, p.4). I believe, that although the critique is valid and supported by previous empirical research on memory and decision processes, the authors fail to present data which rule out dynamic and early effects of language on "low-level" processes.

In fact, the data presented in Chapter 5 informs this debate since it reports a modulation of visual processing by labels. This effect appeared completely

independently of participants' memory and/or decision processes since it arises for unattended stimuli, thus supporting interpretation provided by Lupyan & Spivey (2008). Furthermore, this observation reinforces the value of investigating Whorfian effects using ERPs since they provide a more accurate picture of the timing and processes influenced by linguistic labels. Behavioural data are probably insufficient to draw firm conclusions because they are subject to influences from decisional and strategic processes. All Klemfuss and colleagues (2012) show that by reducing or cutting down the involvement of such processes, behavioural effects are lost.

9. Limitations and avenues for further study

Although the results presented in this thesis have provided evidence in favour of an influence of language on other cognitive processes, they represent early steps in our understanding of the interconnections between linguistic, categorical, and perceptual processes at a neural level. Indeed, compared to most of the behavioural evidence in the field, the studies reported here have only investigated and reported basic effects of language but have not looked into conditions where access to language is disrupted. The use of verbal interference tasks, for instance, has so far been restricted to behavioural studies. In a way this is not surprising, since, from a technical point of view, such tasks require the constant production of speech which would lead to major EEG artefacts produced by articulatory muscles. However, changes in the type of interference task (e.g., silent rehearsing of rote series, report of a heard sequence, mental counting) could potentially approximate the previously used conditions of

verbal interference. Being able to implement such conditions while recording neurophysiological data would be interesting given the higher sensitivity offered by ERP measures. Indeed, the consensus emerging from behavioural studies using verbal interference is that, under such conditions, Whorfian effects seem to disappear. However, the data presented in this thesis suggest that RT and accuracy measures are unaffected even without resorting to verbal interference. This begs the question of whether the absence of linguistic relativity effects in conditions of verbal interference during behavioural tasks necessarily imply a lack of such effects in ERPs. Only after such an issue is cleared can researchers start formulating a clear model of how and when Whorfian effects arise. If verbal interference does indeed cancel the influence of language on categorical and perceptual processes, this must be explained theoretically. The label-feedback hypothesis does allow for the possibility that when a participant is required to perform a verbal task while categorising or perceiving stimuli, the influence of language may be reduced, if not cancelled. More importantly, even though verbal interference is not often questioned in the literature, it remains unclear what exactly is being disrupted by verbal interference and whether the effect is indeed limited to language. Such questioning is not far-fetched since participants are essentially performing a dual-task in which cognitive processes such as inhibition and control are likely at play (e.g., stroop-type tasks). Without more accurate data on such conditions it seems almost impossible to reach any definitive conclusions from studies having used verbal interference. Future studies will therefore need to investigate such a possibility using appropriate methods and control conditions.

The early effects of lexical labels on visual processing reported in the case of cups and mugs in conjunction with behavioural data such as that provided by Gilbert *et al.* (2008) or Lupyan and Spivey (2008; 2010a) provide strong evidence for an influence of language on early visual processing. However, in all those cases the items under scrutiny were perceptually and conceptually related: cups and mugs, cats and dogs, letters or numbers. The influence of label on the discrimination of those closely related items may be particularly beneficial but one could wonder whether this would be the case in the opposite situation where the items are strongly dissimilar but the lexical labels are the same, as is the case for polysemy. In other words, would speakers of English consider a financial establishment conceptually related to “riverside” given the fact that these two concepts are both labelled “bank”, an effect that would arguably not apply to speakers of French who have different labels for the two objects? In that situation, the visual system would not benefit from the activation of labels in order to discriminate perceptually distant stimuli but label effects may still arise. Such result would shed light on the amount of functional involvement of language in basic visual processes, whereby label influence may either apply “blindly” across the board or be called upon, depending on its relevance and beneficial effect. Such an aspect will therefore need to be implemented in models accounting for Whorfian effects.

In sum, only the tip of the iceberg has been uncovered. While it is clear that language does have an influence on other processes, little is known about the source and nature of these effects especially regarding their underlying mechanisms. Although relatively new in the field of EEG research, analyses of

functional cortical connectivity may help shed light on the nature of interactions between language and thought. Furthermore, given the varied levels where language has been shown to have an effect, it would be valuable to determine the representational levels and magnitude of linguistic relativity effects as well as their interactions. For instance, what types of linguistic features will elicit a Whorfian effect? What representation levels will be affected by them? Is it that some features (labels vs morpho-syntactic features) are more relevant than others and therefore impact cognitive processes such as categorisation and perception differently?

10. Concluding Remarks

1. Grammatical features such as gender can influence speakers' object categorisation even when task-irrelevant.
2. Terminology effects on visual processing extend beyond the domain of colour perception and modulate speakers' perception of real objects such as cups and mugs.
3. Links in the lexicon that are idiosyncratic to language where two unrelated concepts are linked in the lexicon via a compound, are measurable on a semantic level (for instance, words like *turtle* and *neck* have become related in semantic/conceptual memory by virtue of their being linked in the compound word *turtleneck*).
4. Whorfian effects can occur at different processing levels (perceptual,

categorical, conceptual).

5. Whorfian effects can appear as strong or weak depending on the experimental requirements of tasks and stimuli.
6. Whorfian effects are perhaps best conceptualised in terms of online top-down modulations of visual or categorical processes rather than being squarely deterministic in nature.
7. Language is not modular nor informationally encapsulated, rather it interacts highly with other cognitive processes, leading to dynamic integration of information in the human brain.

References

- Ameel, E., Malt, B. C., Storms, G., & Van Assche, F. (2009). Semantic convergence in the bilingual lexicon. *Journal of Memory and Language*, 60(2), 270–290. doi:10.1016/j.jml.2008.10.001
- Ameel, E., Storms, G., Malt, B. C., & Sloman, S. A. (2005). How bilinguals solve the naming problem. *Journal of Memory and Language*, 53(1), 60–80. doi:10.1016/j.jml.2005.02.004
- Andrews, S., Miller, B., & Rayner, K. (2004). Eye movements and morphological segmentation of compound words: There is a mouse in mousetrap. *European Journal of Cognitive Psychology*, 16(1), 285–311. doi:10.1080/09541440340000123
- Aronoff, M. (1994). *Morphology by itself: stems and inflectional classes*. Cambridge, Mass.: MIT Press.
- Athanasopoulos, P. (2006). Effects of the grammatical representation of number on cognition in bilinguals. *Bilingualism: Language and Cognition*, 9(01), 89. doi:10.1017/S1366728905002397
- Athanasopoulos, P. (2007). Interaction between grammatical categories and cognition in bilinguals: The role of proficiency, cultural immersion, and language of instruction. *Language and Cognitive Processes*, 22(5), 689–699. doi:10.1080/01690960601049347
- Athanasopoulos, P. (2009). Cognitive representation of colour in bilinguals: The case of Greek blues. *Bilingualism: Language and Cognition*, 12(01), 83. doi:10.1017/S136672890800388X

- Athanasopoulos, P., & Bylund, E. (2012). Does Grammatical Aspect Affect Motion Event Cognition? A Cross-Linguistic Comparison of English and Swedish Speakers. *Cognitive science*, 37(2), 286–309. doi:10.1111/cogs.12006
- Athanasopoulos, P., & Kasai, C. (2008). Language and thought in bilinguals: The case of grammatical number and nonverbal classification preferences. *Applied Psycholinguistics*, 29(01). doi:10.1017/S0142716408080053
- Athanasopoulos, P., Damajanovic, L., Krajciova, A., & Sasakai, M. (2010a). Representation of colour concepts in bilingual cognition: The case of Japanese blues. *Bilingualism: Language and Cognition*, 14(01), 9–17. doi:10.1017/S1366728909990046
- Athanasopoulos, P., Dering, B., Wiggett, A., Kuipers, J.-R., & Thierry, G. (2010b). Perceptual shift in bilingualism: Brain potentials reveal plasticity in pre-attentive colour perception. *Cognition*, 116(3), 437–443. doi:10.1016/j.cognition.2010.05.016
- Balaban, M. T., & Waxman, S. R. (1997). Do words facilitate object categorization in 9-month-old infants? *Journal of Experimental Child Psychology*, 64(1), 3–26. doi:10.1006/jecp.1996.2332
- Baldwin, D. A. D., & Markman, E. M. E. (1989). Establishing word-object relations: a first step. *Child Development*, 60(2), 381–398. doi:10.2307/1130984
- Barsalou, L. W. (1992). *Cognitive psychology: an overview for cognitive scientists*. Hillsdale, N.J.: L. Erlbaum Associates.

- Barsalou, L. W., (2008). Grounded Cognition. *The Annual Review of Psychology*, 59, 617–645. doi:10.1146/annurev.psych.59.103006.093639
- Bassetti, B. (2007). Bilingualism and thought: Grammatical gender and concepts of objects in Italian-German bilingual children. *International Journal of Bilingualism*, 11(3), 251–273. doi:10.1177/13670069070110030101
- Behl-Chadha, G. G. (1996). Basic-level and superordinate-like categorical representations in early infancy. *Cognition*, 60(2), 105–141. doi: 10.1016/0010-0277(96)00706-8
- Bentin, S., Kutas, M., & Hillyard, S. A. (1993). Electrophysiological evidence for task effects on semantic priming in auditory word processing. *Psychophysiology*, 30(2), 161–169. doi:10.1111/j.1469-8986.1993.tb01729.x
- Bentin, S., McCarthy, G., & Wood, C. C. (1985). Event-related potentials, lexical decision and semantic priming. *Electroencephalography and clinical Neurophysiology*, 60(4), 343–355.
- Berlin, B., & Kay, P. (1969). *Basic color terms; their universality and evolution*. Berkeley: University of California Press.
- Blackford, T., Holcomb, P. J., Grainger, J., & Kuperberg, G. R. (2012). A funny thing happened on the way to articulation: N400 attenuation despite behavioral interference in picture naming. *Cognition*, 123(1), 84–99. doi: 10.1016/j.cognition.2011.12.007

- Borges, J. L. (1962). *Funes el memorioso*. Retrieved from <http://www.bridgewater.edu/~atrupe/GEC101/Funes.html>
- Borghi, A. M., & Pecher, D. (2011). Introduction to the Special Topic Embodied and Grounded Cognition. *Frontiers in psychology, 2*. doi:10.3389/fpsyg.2011.00187
- Boroditsky, L. L., & Schmidt, L. A. (2000). Sex, Syntax, and Semantics. *Proceedings of the Twenty-Second Annual Meeting of the Cognitive Science Society*, 1–6.
- Boroditsky, L., Schmidt, L. A., & Phillips, W. (2003). Sex, syntax, and semantics. *Language in mind: Advances in the study of language and cognition*, 61–79.
- Boutonnet, B., & Athanasopoulos, P. (unpublished manuscript). Learning grammatical gender in a second language changes categorization of inanimate objects: Replications and new evidence from English learners of L2 French.
- Boutonnet, B., Athanasopoulos, P., & Thierry, G. (2012). Unconscious effects of grammatical gender during object categorisation. *Brain Research, 1479*, 72–79. doi:10.1016/j.brainres.2012.08.044
- Bowermann, M., & Choi, S. (1991). Learning to express motion events in English and Korean: The influence of language-specific lexicalization patterns. *Cognition, 41*, 83–121.

- Cangelosi, A., & Harnad, S. (2001). The adaptive advantage of symbolic theft over sensorimotor toil: Grounding language in perceptual categories. *Evolution of communication*, 4(1), 117–142.
- Casasola, M. (2005). Can Language Do the Driving? The Effect of Linguistic Input on Infants' Categorization of Support Spatial Relations. *Developmental Psychology*, 41(1), 183–192. doi: 10.1037/0012-1649.41.1.183
- Chomsky, N. (1965). *Aspects of the theory of syntax*. Cambridge: M.I.T. Press.
- Chomsky, N. (2000). *New horizons in the study of language and mind*. Cambridge [England]; New York: Cambridge University Press.
- Churchland, P. S., Ramachandran, V. S., & Sejnowski, T. J. (1994). A Critique of Pure Vision. In C. Koch & J. Davis (Eds.), *Large-Scale Neuronal Theories of the Brain* (pp. 23–60). Cambridge, MA: MIT Press.
- Clarke, M. A., Losoff, A., Dickenson, M., McCracken, M. D., & Still, J. (1981). Gender perception in Arabic and English. *Language Learning*, 31(1), 159–169.
- Clifford, A., Franklin, A., Davies, I. R. L., & Holmes, A. (2009). Electrophysiological markers of categorical perception of color in 7-month old infants. *Brain and Cognition*, 71(2), 165–172. doi:10.1016/j.bandc.2009.05.002

- Clifford, A., Holmes, A., Davies, I. R. L., & Franklin, A. (2010). Color categories affect pre-attentive color perception. *Biological Psychology*, 85(2), 275–282. doi:10.1016/j.biopsycho.2010.07.014
- Cook, V., Bassetti, B., Kasai, C., Sasaki, M., & Takahashi, J. A. (2006). Do bilinguals have different concepts? The case of shape and material in Japanese L2 users of English. *International Journal of Bilingualism*, 10(2), 137–152. doi:10.1177/13670069060100020201
- Corbett, G. G. (1991). *Gender*. Cambridge [England]; New York: Cambridge University Press.
- Costa, A., Strijkers, K., Martin, C., & Thierry, G. (2009). The time course of word retrieval revealed by event-related brain potentials during overt speech. *Proceedings of the National Academy of Sciences of the United States of America*, 106(50), 21442–21446. doi:10.1073/pnas.0908921106
- Coulson, S., King, J. W., & Kutas, M. (1998). Expect the Unexpected: Event-related Brain Response to Morphosyntactic Violations. *Language and Cognitive Processes*, 13(1), 21–58. doi:10.1080/016909698386582
- Csibra, G., Czigler, I., & Ambro, A. (1994). Effects of stimulus alternation, repetition and response requirements on event-related potentials to patterned visual stimuli. *Biological Psychology*, 37(2), 115–132.
- Cubelli, R., Lotto, L., Paolieri, D., Girelli, M., & Job, R. (2005). Grammatical gender is selected in bare noun production: Evidence from the picture–word interference paradigm. *Journal of Memory and Language*, 53(1), 42–59. doi:10.1016/j.jml.2005.02.007

- Cubelli, R., Paolieri, D., Lotto, L., & Job, R. (2011). The effect of grammatical gender on object categorization. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 37(2), 449–460. doi:10.1037/a0021965
- Czigler, I., Balázs, L., & Pató, L. G. (2004). Visual change detection: event-related potentials are dependent on stimulus location in humans. *Neuroscience Letters*, 364(3), 149–153. doi:10.1016/j.neulet.2004.04.048
- Czigler, I., Balázs, L., & Winkler, I. (2002). Memory-based detection of task-irrelevant visual changes. *Psychophysiology*, 39(6), 869–873.
- Daoutis, C. A., Franklin, A., Riddett, A., Clifford, A., & Davies, I. R. L. (2010). Categorical effects in children's colour search: A cross-linguistic comparison. *British Journal of Developmental Psychology*, 24(2), 373–400. doi:10.1348/026151005X51266
- Davidoff, J. J., Davies, I. I., & Roberson, D. D. (1999). Colour categories in a stone-age tribe. *Nature*, 398(6724), 203–204. doi:10.1038/18335
- Davies, I. R., & Corbett, G. G. (1997). A cross-cultural study of colour grouping: evidence for weak linguistic relativity. *British Journal of Psychology*, 88 (Pt 3), 493–517.
- Dering, B., Martin, C. D., Moro, S., Pegna, A. J., & Thierry, G. (2011). Face-sensitive processes one hundred milliseconds after picture onset. *Frontiers in human neuroscience*, 5, 93–93. doi:10.3389/fnhum.2011.00093

- Desimone, R., Albright, T. D., Gross, C. G., & Bruce, C. (1984). Stimulus-selective properties of inferior temporal neurons in the macaque. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, 4(8), 2051–2062.
- Dölen, G., Osterweil, E., Rao, B. S. S., Smith, G. B., Auerbach, B. D., Chattarji, S., & Bear, M. F. (2007). Correction of Fragile X Syndrome in Mice. *Neuron*, 56(6), 955–962. doi:10.1016/j.neuron.2007.12.001
- Drivonikou, G. V., Kay, P., Regier, T., Ivry, R. B., Gilbert, A. L., Franklin, A., & Davies, I. R. L. (2007). Further evidence that Whorfian effects are stronger in the right visual field than the left. *Proceedings of the National Academy of Sciences*, 104(3), 1097–1102. doi:10.1073/pnas.0610132104
- Duñabeitia, J. A., Laka, I., Perea, M., & Carreiras, M. (2008). Is *Milkman* a superhero like *Batman*? Constituent morphological priming in compound words. *European Journal of Cognitive Psychology*, 21(4), 615–640. doi:10.1080/09541440802079835
- Eimas, P. D., & Quinn, P. C. (2008). Studies on the Formation of Perceptually Based Basic-Level Categories in Young Infants. *Child Development*, 65(3), 903–917.
- Elman, J. L. (2004). An alternative view of the mental lexicon. *Trends in Cognitive Sciences*, 8(7), 301–306. doi:10.1016/j.tics.2004.05.003
- Elston-Guttler, K., Gunter, T., & Kotz, S. (2005). Zooming into L2: Global language context and adjustment affect processing of interlingual

- homographs in sentences. *Cognitive Brain Research*, 25(1), 57–70. doi: 10.1016/j.cogbrainres.2005.04.007
- Feigenson, L., Dehaene, S., & Spelke, E. (2004). Core systems of number. *Trends in Cognitive Sciences*, 8(7), 307–314. doi:10.1016/j.tics.2004.05.002
- Fiorentino, R., & Poeppel, D. (2007). Compound words and structure in the lexicon. *Language and Cognitive Processes*, 22(7), 953–1000. doi: 10.1080/01690960701190215
- Flaherty, M. (2001). How a Language Gender System Creeps into Perception. *Journal of Cross-Cultural Psychology*, 32(1), 18–31. doi: 10.1177/0022022101032001005
- Flecken, M. (2010). Event conceptualization by early Dutch–German bilinguals: Insights from linguistic and eye-tracking data. *Bilingualism: Language and Cognition*, 14(01), 61–77. doi:10.1017/S1366728910000027
- Fodor, J. A. (1975). *The language of thought*. Cambridge, Mass.: Harvard University Press.
- Fodor, J. A. (2008). *LOT 2 : the language of thought revisited*. Oxford; New York: Clarendon Press ; Oxford University Press.
- Folstein, J. R., & Van Petten, C. (2007). Influence of cognitive control and mismatch on the N2 component of the ERP: A review. *Psychophysiology*, 0(0), 070915195953001–??? doi:10.1111/j.1469-8986.2007.00602.x
- Fonteneau, E., & Davidoff, J. (2007). Neural correlates of colour categories. *Neuroreport*, 18(13), 1323–1327. doi: 10.1097/WNR.0b013e3282c48c33

- Forbes, J. N., Poulin-Dubois, D., Rivero, M. R., & Sera, M. D. (2008). Grammatical Gender Affects Bilinguals' Conceptual Gender: Implications for Linguistic Relativity and Decision Making. *Open Applied Linguistics Journal*, 1, 68–76.
- Forgács, B., Bohrn, I., Baudewig, J., Hofmann, M. J., Pléh, C., & Jacobs, A. M. (2012). Neural correlates of combinatorial semantic processing of literal and figurative noun noun compound words. *NeuroImage*, 1–11. doi: 10.1016/j.neuroimage.2012.07.029
- Foxe, J., & Simpson, G. (2002). Flow of activation from V1 to frontal cortex in humans. *Experimental Brain Research*, 142(1), 139–150. doi:10.1007/s00221-001-0906-7
- Franklin, A., Drivonikou, G., Clifford, A., Kay, P., Regier, T., & Davies, I. (2008). Lateralization of categorical perception of color changes with color term acquisition. *Proceedings of the National Academy of Sciences*, 105(47), 18221.
- Freeman, W. J. (2000). *How brains make up their minds*. New York: Columbia University Press.
- Friederici. (1995). The Time Course of Syntactic Activation During Language Processing: A Model Based on Neuropsychological and Neurophysiological Data. *Brain and Language*, 50(3), 23–23. doi: 10.1006/brln.1995.1048

- Friederici, A. D., & Jacobsen, T. (1999). Processing grammatical gender during language comprehension. *Journal of psycholinguistic research*, 28(5), 467–484. doi:10.1006/jmla.1999.2649
- Friederici, A. D., Pfeifer, E., & Hahne, A. (1993). Event-related brain potentials during natural speech processing: effects of semantic, morphological and syntactic violations. *Brain research. Cognitive brain research*, 1(3), 183–192.
- Friederici, A. D., Steinhauer, K., & Frisch, S. (1999). Lexical integration: sequential effects of syntactic and semantic information. *Memory & cognition*, 27(3), 438–453.
- Gagné, C. L., & Spalding, T. L. (2009). Constituent integration during the processing of compound words: Does it involve the use of relational structures? *Journal of Memory and Language*, 60(1), 20–35. doi:10.1016/j.jml.2008.07.003
- Gallistel, C. R. (1989). Animal cognition: the representation of space, time and number. *Psychology*, 40, 155–189. doi:10.1146/annurev.ps.40.020189.001103
- Ganis, G., Kutas, M., & Sereno, M. I. (1996). The search for “common sense”: An electrophysiological study of the comprehension of words and pictures in reading. *Journal of cognitive neuroscience*, 8(2), 89–106.
- Ganushchak, L. Y., Verdonschot, R. G., & Schiller, N. O. (2011). When leaf becomes neuter. *Neuroreport*, 22(3), 106–110. doi:10.1097/WNR.0b013e3283427359

- Gentner, D., & Goldin-Meadow, S. (2003). *Language in mind : advances in the study of language and thought*. Cambridge, Mass.: MIT Press.
- Gilbert, A. L., Regier, T., Kay, P., & Ivry, R. B. (2006). Whorf hypothesis is supported in the right visual field but not the left. *Proceedings of the National Academy of Sciences*, *103*(2), 489–494. doi:10.2307/30048329
- Gilbert, A., Regier, T., Kay, P., & Ivry, R. (2008). Support for lateralization of the Whorf effect beyond the realm of color discrimination. *Brain and Language*, *105*(2), 91–98. doi:10.1016/j.bandl.2007.06.001
- Gilbert, C. D., & Sigman, M. (2007). Brain States: Top-Down Influences in Sensory Processing. *Neuron*, *54*(5), 677–696. doi:10.1016/j.neuron.2007.05.019
- Gordon, P. (2004). Numerical Cognition Without Words: Evidence from Amazonia. *Science*, *306*(5695), 496–499. doi:10.1126/science.1094492
- Gratton, G., Coles, M. G. H., & Donchin, E. (1983). A new method for off-line removal of ocular artifact. *Electroencephalography and clinical Neurophysiology*, *55*(4), 468–484.
- Graves, W. W., Binder, J. R., Desai, R. H., Conant, L. L., & Seidenberg, M. S. (2010). Neural correlates of implicit and explicit combinatorial semantic processing. *NeuroImage*, *53*(2), 638–646. doi:10.1016/j.neuroimage.2010.06.055
- Grosjean, F. (1998). Transfer and language mode. *Bilingualism: Language and Cognition*, *1*(03), 175–176.

- Gumperz, J. J., & Levinson, S. C. (1996). *Rethinking linguistic relativity*.
Cambridge; New York, NY, USA: Cambridge University Press.
- Gunter, T. C., Friederici, A. D., & Schriefers, H. (2000). Syntactic gender and semantic expectancy: ERPs reveal early autonomy and late interaction. *Journal of cognitive neuroscience*, *12*(4), 556–568.
- Guthrie, D., & Buchwald, J. S. (1991). Significance testing of difference potentials. *Psychophysiology*, *28*(2), 240–244.
- Hahne, A. A., & Friederici, A. D. A. (2002). Differential task effects on semantic and syntactic processes as revealed by ERPs. *Brain research. Cognitive brain research*, *13*(3), 339–356. doi:10.1016/S0926-6410(01)00127-6
- Hahne, A., & Friederici, A. D. (1999). Electrophysiological evidence for two steps in syntactic analysis. Early automatic and late controlled processes. *Journal of cognitive neuroscience*, *11*(2), 194–205.
- Hare, B., Call, J., & Tomasello, M. (2001). Do chimpanzees know what conspecifics know? *Animal Behaviour*, *61*(1), 139–151. doi:10.1006/anbe.2000.1518
- Harnad, S. (2005). To Cognize is to Categorize: Cognition is Categorization. In C. Lefebvre & H. Cohen (Eds.), *Handbook on Categorization* (pp. 1–29). Elsevier.
- Holcomb, P. J. P., & McPherson, W. B. W. (1994). Event-related brain potentials reflect semantic priming in an object decision task. *Brain and Cognition*, *24*(2), 259–276. doi:10.1006/brcg.1994.1014

- Holcomb, P. J., & Neville, H. J. (1990). Auditory and Visual Semantic Priming in Lexical Decision: A Comparison Using Event-related Brain Potentials. *Language and Cognitive Processes*, 5(4), 281–312. doi: 10.1080/01690969008407065
- Hoover, J. E., & Strick, P. L. (1993). Multiple output channels in the basal ganglia. *Science*, 259(5096), 819–821.
- Humphreys, G. W., Riddoch, M. J., & Price, C. J. (1997). Top-down processes in object identification: evidence from experimental psychology, neuropsychology and functional anatomy. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 352(1358), 1275–1282. doi: 10.1098/rstb.1997.0110
- Hunt, E., & Agnoli, F. (1991). The Whorfian hypothesis: A cognitive psychology perspective. *Psychological Review*, 98(3), 377.
- Hyönä, J., Bertram, R., & Pollatsek, A. (2004). Are long compound words identified serially via their constituents? Evidence from an eye-movement-contingent display change study. *Memory & cognition*, 32(4), 523–532.
- Imai, M. M., & Gentner, D. D. (1997). A cross-linguistic study of early word meaning: universal ontology and linguistic influence. *Cognition*, 62(2), 169–200. doi:10.1016/S0010-0277(96)00784-6
- James, W. (1890). *Principles of psychology*. New York: H. Holt and Co.

- Jiang, X., Bradley, E., Rini, R. A., Zeffiro, T., VanMeter, J., & Riesenhuber, M. (2007). Categorization training results in shape-and category-selective human neural plasticity. *Neuron*, 53(6), 891.
- Jordan, G., & Mollon, J. D. (1997). Adaptation of colour vision to sunlight. *Nature*, 386(6621), 135–136. doi:10.1038/386135b0
- Kelly, R. M., & Strick, P. L. (2004). Macro-architecture of basal ganglia loops with the cerebral cortex: use of rabies virus to reveal multisynaptic circuits. *Prog Brain Res, Volume 143*, 447–459.
- Kiefer, M., & Pulvermüller, F. (2012). Conceptual representations in mind and brain: Theoretical developments, current evidence and future directions. *Cortex*, 48(7), 805–825. doi:10.1016/j.cortex.2011.04.006
- Kiss, G. R., Armstrong, C., Milroy, R., & Piper, J. (1973). An associative thesaurus of English and its computer analysis. In A. J. Aitken, R. W. Bailey, & N. Hamilton-Smith (Eds.), *The Computer and Literary Studies*. Edinburgh: Edinburgh University Press.
- Klemfuss, N., Prinzmetal, W., & Ivry, R. B. (2012). How does language change perception: a cautionary note. *Frontiers in psychology*, 3, 78–78. doi:10.3389/fpsyg.2012.00078/abstract
- Koester, D., Gunter, T. C., & Wagner, S. (2007). The morphosyntactic decomposition and semantic composition of German compound words investigated by ERPs. *Brain and Language*, 102(1), 64–79. doi:10.1016/j.bandl.2006.09.003

- Koivisto, M., Railo, H., Revonsuo, A., Vanni, S., & Salminen-Vaparanta, N. (2011). Recurrent Processing in V1/V2 Contributes to Categorization of Natural Scenes. *Journal of Neuroscience*, *31*(7), 2488–2492. doi:10.1523/JNEUROSCI.3074-10.2011
- Kolk, H. H. J., Chwilla, D. J., van Herten, M., & Oor, P. J. W. (2003). Structure and limited capacity in verbal working memory: A study with event-related potentials. *Brain and Language*, *85*(1), 1–36. doi:10.1016/S0093-934X(02)00548-5
- Kolk, H., & Chwilla, D. (2007). Late positivities in unusual situations. *Brain and Language*, *100*(3), 257–261. doi:10.1016/j.bandl.2006.07.006
- Konishi, T. (1994). The connotations of gender: A semantic differential study of German and Spanish. *Word*, (45), 317–327.
- Kousta, S.-T., Vinson, D. P., & Vigliocco, G. (2008). Investigating linguistic relativity through bilingualism: The case of grammatical gender. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *34*(4), 843–858. doi:10.1037/0278-7393.34.4.843
- Kuipers, J.-R., & Thierry, G. (2011). N400 amplitude reduction correlates with an increase in pupil size. *Frontiers in human neuroscience*, *5*, 61. doi:10.3389/fnhum.2011.00061
- Kurinski, E., & Sera, M. D. (2010). Does learning Spanish grammatical gender change English-speaking adults' categorization of inanimate objects? *Bilingualism: Language and Cognition*, *14*(02), 203–220. doi:10.1017/S1366728910000179

- Kutas, M., & Federmeier, K. D. (2011). Thirty Years and Counting: Finding Meaning in the N400 Component of the Event-Related Brain Potential (ERP). *Annual review of psychology*, 62(1), 621–647. doi:10.1146/annurev.psych.093008.131123
- Kutas, M., & Hillyard, S. A. (1980). Reading senseless sentences: Brain potentials reflect semantic incongruity. *Science*, 207, 203–205.
- Kutas, M., & Hillyard, S. A. (1984). Brain potentials during reading reflect word expectancy and semantic association. *Nature*, 161–163.
- Lamme, V. A., & Roelfsema, P. R. (2000). The distinct modes of vision offered by feedforward and recurrent processing. *Trends in neurosciences*, 23(11), 571–579.
- Lee, P. (1996). *The whorf theory complex: A critical reconstruction*. Amsterdam; Philadelphia: J. Benjamins Pub.
- Lemhöfer, K., Spalek, K., & Schriefers, H. (2008). Cross-language effects of grammatical gender in bilingual word recognition and production. *Journal of Memory and Language*, 59(3), 312–330. doi:10.1016/j.jml.2008.06.005
- Li, S., Mayhew, S. D., & Kourtzi, Z. (2009). Learning Shapes the Representation of Behavioral Choice in the Human Brain. *Neuron*, 62(3), 441–452. doi:10.1016/j.neuron.2009.03.016
- Liu, C., Tardif, T., Mai, X., Gehring, W. J., Simms, N., & Luo, Y.-J. (2010a). What's in a name? Brain activity reveals categorization processes differ

across languages. *Human Brain Mapping*, NA–NA. doi:10.1002/hbm.20974

Liu, Q., Li, H., Campos, J. L., Teeter, C., Tao, W., Zhang, Q., & Sun, H.-J. (2010b). Language suppression effects on the categorical perception of colour as evidenced through ERPs. *Biological Psychology*, 85(1), 45–52. doi:10.1016/j.biopsycho.2010.05.001

Logothetis, N. K., & Sheinberg, D. L. (1996). Visual object recognition. *Annual review of neuroscience*, 19, 577–621. doi:10.1146/annurev.ne.19.030196.003045

Love, B. C., Medin, D. L., & Gureckis, T. M. (2004). SUSTAIN: A Network Model of Category Learning. *Psychological Review*, 111(2), 309–332. doi:10.1037/0033-295X.111.2.309

Lu, A., Hodges, B. H., Zhang, J., & Wang, X. (2012). A Whorfian speed bump? Effects of Chinese color names on recognition across hemispheres. *Language Sciences*, 34(5), 591–603. doi:10.1016/j.langsci.2012.03.014

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

Lucy, J. A. (1992a). *Language Diversity and Thought: A Reformulation of the Linguistic Relativity Hypothesis*. Cambridge Univ Pr.

Lucy, J. A. (1992b). *Grammatical categories and cognition : a case study of the linguistic relativity hypothesis*. Cambridge [England]; New York, NY, USA: Cambridge University Press.

- Lucy, J. A. (1997). Linguistic Relativity. *Annual Review of Anthropology*, 26, 291–312.
- Lupyan, G. (2012). Linguistically modulated perception and cognition: the label-feedback hypothesis. *Frontiers in psychology*, 3, 54–54. doi:10.3389/fpsyg.2012.00054
- Lupyan, G., & Spivey, M. J. (2008). Perceptual processing is facilitated by ascribing meaning to novel stimuli. *Current Biology*, 18(10), 0–0. doi:10.1016/j.cub.2008.02.073
- Lupyan, G., & Spivey, M. J. (2010a). Redundant spoken labels facilitate perception of multiple items. *Attention, Perception, & Psychophysics*, 72(8), 2236–2253. doi:10.3758/APP.72.8.2236
- Lupyan, G., & Spivey, M. J. (2010b). Making the invisible visible: Verbal but not visual cues enhance visual detection. *PLoS ONE*, 5(7), e11452. doi:10.1371/journal.pone.0011452.g001
- 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(12), 1077–1083. doi:10.1111/j.1467-9280.2007.02028.x
- Lupyan, G., Thompson-Schill, S. L., & Swingley, D. (2010). Conceptual Penetration of Visual Processing. *Psychological Science*, 21(5), 682–691. doi:10.1177/0956797610366099

- Luria, A. (1986) *The Mind of a Mnemonist: A Little Book about a Vast Memory*
Cambridge MA: Harvard University Press. .
- Mahon, B. Z., & Caramazza, A. (2008). A critical look at the embodied cognition hypothesis and a new proposal for grounding conceptual content. *Journal of Physiology-Paris*, 102(1-3), 59–70. doi:10.1016/j.jphysparis.2008.03.004
- Mahon, B. Z., & Caramazza, A. (2009). Concepts and Categories: A Cognitive Neuropsychological Perspective. *Annual review of psychology*, 60(1), 27–51. doi:10.1146/annurev.psych.60.110707.163532
- Majid, A., Bowerman, M., Kita, S., Haun, D. B. M., & Levinson, S. C. (2004). Can language restructure cognition? The case for space. *Trends in Cognitive Sciences*, 8(3), 108–114. doi:10.1016/j.tics.2004.01.003
- Martin, A. (2007). The Representation of Object Concepts in the Brain. *Annual review of psychology*, 58(1), 25–45. doi:10.1146/annurev.psych.57.102904.190143
- Martinez, I. M., & Shatz, M. (1996). Linguistic influences on categorization in preschool children: A crosslinguistic study. *Journal of Child Language*, 23, 529–546. doi:10.1017/S030500090000893X
- McClelland, J. L. J., McNaughton, B. L. B., & O'Reilly, R. C. R. (1995). Why there are complementary learning systems in the hippocampus and neocortex: insights from the successes and failures of connectionist models of learning and memory. *Psychological Review*, 102(3), 419–457.

- McClelland, J. L., & Rumelhart, D. E. (1981). An interactive activation model of context effects in letter perception: I. An account of basic findings. *Psychological Review*, 88(5), 375.
- Mesulam, M. M. (1998). From sensation to cognition. *Brain : a journal of neurology*, 121 (Pt 6), 1013–1052.
- Miceli, G. G., Turriziani, P. P., Caltagirone, C. C., Capasso, R. R., Tomaiuolo, F. F., & Caramazza, A. A. (2002). The neural correlates of grammatical gender: an fMRI investigation. *Journal of cognitive neuroscience*, 14(4), 618–628. doi:10.1162/08989290260045855
- Mo, L., Xu, G., Kay, P., & Tan, L. H. (2011). Electrophysiological evidence for the left-lateralized effect of language on preattentive categorical perception of color. *Proceedings of the National Academy of Sciences*, 108(34), 14026–14030. doi:10.1073/pnas.1111860108
- Muhammad, R., Wallis, J. D., & Miller, E. K. (2006). A comparison of abstract rules in the prefrontal cortex, premotor cortex, inferior temporal cortex, and striatum. *Journal of cognitive neuroscience*, 18(6), 974–989. doi: 10.1162/jocn.2006.18.6.974
- Müller, F. M. (1909). *The science of thought*. London: Longmans, Green & Co.
- Näätänen, R. (1995). The mismatch negativity: a powerful tool for cognitive neuroscience. *Ear Hear*, 16(1), 6–18.

- Näätänen, R. (2009). The role of attention in auditory information processing as revealed by event-related potentials and other brain measures of cognitive function, 1–88.
- Oakes, L. M. L., & Madole, K. L. K. (2000). The future of infant categorization research: a process-oriented approach. *Child Development*, 71(1), 119–126. doi:10.2307/1132224
- Osterhout, L., & Holcomb, P. J. (1992). Event-related brain potentials elicited by syntactic anomaly. *Journal of Memory and Language*, 31(6), 785–806.
- Ozgen, E. (2004). Language, Learning, and Color Perception. *Current Directions in Psychological Science*, 13(3), 95–98. doi:10.1111/j.0963-7214.2004.00282.x
- Ozgen, E., & Davies, I. R. (1998). Turkish color terms: tests of Berlin and Kay's theory of color universals and linguistic relativity. *Linguistics*, 36(5), 919–956.
- Papafragou, A., & Selimis, S. (2010). Event categorisation and language: A cross-linguistic study of motion. *Language and Cognitive Processes*, 25(2), 224–260. doi:10.1080/01690960903017000
- Pavlenko, A., & Malt, B. C. (2010). Kitchen Russian: Cross-linguistic differences and first-language object naming by Russian–English bilinguals. *Bilingualism: Language and Cognition*, 14(01), 19–45. doi:10.1017/S136672891000026X

- Pazo-Alvarez, P., Cadaveira, F., & Amenedo, E. (2003). MMN in the visual modality: a review. *Biological Psychology*, 63(3), 199–236. doi:10.1016/S0301-0511(03)00049-8
- Penn, D. C., Holyoak, K. J., & Povinelli, D. J. (2008). Darwin's mistake: Explaining the discontinuity between human and nonhuman minds. *The Behavioral and brain sciences*, 31(02). doi:10.1017/S0140525X08003543
- Phillips, W., & Boroditsky, L. (2003). Can quirks of grammar affect the way you think? Grammatical gender and object concepts. *Proceedings of the 25th annual meeting of the Cognitive Science Society*, 928–933.
- Phillips, W., & Santos, L. R. (2007). Evidence for kind representations in the absence of language: Experiments with rhesus monkeys (*Macaca mulatta*). *Cognition*, 102(3), 455–463. doi:10.1016/j.cognition.2006.01.009
- Picton, T., Bentin, S., Berg, P., Donchin, E., Hillyard, S., Johnson, R., Jr, *et al.* (2000). Guidelines for using human event-related potentials to study cognition: Recording standards and publication criteria. *Psychophysiology*, 37(2), 127–152.
- Pinker, S. (1995). *The language instinct : the new science of language and mind*. London: Penguin.
- Pinker, S. (2007). *The stuff of thought : language as a window into human nature*. New York: Viking.

- Plunkett, K., Hu, J.-F., & Cohen, L. B. (2008). Labels can override perceptual categories in early infancy. *Cognition*, *106*(2), 665–681. doi:10.1016/j.cognition.2007.04.003
- Pruden, S. M., Roseberry, S., Göksun, T., Hirsh-Pasek, K., & Golinkoff, R. M. (2012). Infant Categorization of Path Relations During Dynamic Events. *Child Development*, *84*(1), 331–345. doi:10.1111/j.1467-8624.2012.01843.x
- Pullum, G. K. (1989). The great Eskimo vocabulary hoax. *Natural language and linguistic theory*, *7*(2), 275–281.
- Pulvermüller, F. (1996). Hebb's concept of cell assemblies and the psychophysiology of word processing. *Psychophysiology*, *33*(4), 317–333. doi:10.1111/j.1469-8986.1996.tb01057.x
- Pulvermüller, F. (1999). Words in the brain's language. *The Behavioral and brain sciences*, *22*, 253–279.
- Pulvermüller, F. (2012). Meaning and the brain: The neurosemantics of referential, interactive, and combinatorial knowledge. *Journal of Neurolinguistics*, *25*(5), 423–459. doi:10.1016/j.jneuroling.2011.03.004
- Ramos, S., & Roberson, D. (2010). What constrains grammatical gender effects on semantic judgements? Evidence from Portuguese. *European Journal of Cognitive Psychology*, 1–16. doi:10.1080/09541441003662734

Regier, T., & Kay, P. (2009). Language, thought, and color: Whorf was half right. *Trends in Cognitive Sciences*, 13(10), 439–446. doi:10.1016/j.tics.

2009.07.001

Roberson, D., Davidoff, J., Davies, I. R. L., & Shapiro, L. R. (2004). The development of color categories in two languages: a longitudinal study.

Journal of Experimental Psychology: General, 133(4), 554–571. doi:

10.1037/0096-3445.133.4.554

Roberson, D., Davidoff, J., Davies, I. R. L., & Shapiro, L. R. (2005). Color categories: Evidence for the cultural relativity hypothesis. *Cognitive Psychology*, 50(4), 378–411. doi:10.1016/j.cogpsych.2004.10.001

Roberson, D., Pak, H., & Hanley, J. R. (2008). Categorical perception of colour in the left and right visual field is verbally mediated: Evidence from Korean. *Cognition*, 107(2), 752–762. doi:10.1016/j.cognition.2007.09.001

Rugg, M. D., Doyle, M. C., & Wells, T. (1995). Word and nonword repetition within-and across-modality: An event-related potential study. *Journal of cognitive neuroscience*, 7(2), 209–227.

Rumelhart, D. E. (1979). Some problems with the notion that words have literal meanings. In A. Ortony (Ed.), *Metaphor and thought* (pp. 71–82).

Cambridge; New York: Cambridge University Press.

Rumelhart, D. E., McClelland, J. L., & The PDP Research Group. (1986). *Parallel Distributed Processing: Explorations in the microstructure of cognition*.

Cambridge, Mass.: MIT Press.

- Saalbach, H., & Imai, M. (2007). Scope of linguistic influence: Does a classifier system alter object concepts? *Journal of Experimental Psychology: General*, 136(3), 485–501. doi:10.1037/0096-3445.136.3.485
- Seeck, M., Michel, C. M., Blanke, O., Thut, G., Landis, T., & Schomer, D. L. (2001). Intracranial Neurophysiological Correlates Related to the Processing of Faces. *Epilepsy & Behavior*, 2(6), 545–557. doi:10.1006/ebeh.2001.0266
- Seeck, M., Michel, C. M., Mainwaring, N., Cosgrove, R., Blume, H., Ives, J., *et al.* (1997). Evidence for rapid face recognition from human scalp and intracranial electrodes. *Neuroreport*, 8(12), 2749–2754.
- Seger, C. A. (2008). How do the basal ganglia contribute to categorization? Their roles in generalization, response selection, and learning via feedback. *Neuroscience & Biobehavioral Reviews*, 32(2), 265–278. doi:10.1016/j.neubiorev.2007.07.010
- Seger, C. A., & Miller, E. K. (2010). Category Learning in the Brain. *Annual review of neuroscience*, 33(1), 203–219. doi:10.1146/annurev.neuro.051508.135546
- Sera, M. D., Berge, C. A., & Castillo-Pintado, J. D. (1994). Grammatical and conceptual forces in the attribution of gender by English and Spanish speakers. *Cognitive Development*, 9(3), 261–292.
- Sera, M. D., Elieff, C., Forbes, J., Burch, M. C., Rodríguez, W., & Dubois, D. P. (2002). When language affects cognition and when it does not: An analysis of grammatical gender and classification. *Journal of*

Experimental Psychology: General, 131(3), 377–397. doi:

10.1037//0096-3445.131.3.377

Slobin, D. I. (1996). From "thought and language" to "thinking for speaking." In J. J. Gumperz & S. C. Levinson (Eds.), *Rethinking Linguistic Relativity*. Cambridge: Cambridge University Press.

Slobin, D. I. (2003). Language and thought online: Cognitive consequences of linguistic relativity. *Language in mind: Advances in the study of language and thought*, 157, 178.

Snodgrass, J. G., & Vanderwart, M. (1980). A standardized set of 260 pictures: Norms for name agreement, image agreement, familiarity, and visual complexity. *Journal of experimental psychology: Human learning and memory*, 6(2), 174.

Steinhauer, K., & Drury, J. E. (2012). On the early left-anterior negativity (ELAN) in syntax studies. *Brain and Language*, 120(2), 135–162. doi: 10.1016/j.bandl.2011.07.001

Strijkers, K., Costa, A., & Thierry, G. (2010). Tracking Lexical Access in Speech Production: Electrophysiological Correlates of Word Frequency and Cognate Effects. *Cerebral Cortex*, 20(4), 912–928. doi:10.1093/cercor/bhp153

Strijkers, K., Holcomb, P. J., & Costa, A. (2011). Conscious intention to speak proactively facilitates lexical access during overt object naming. *Journal of Memory and Language*, 65(4), 345–362. doi:10.1016/j.jml.2011.06.002

- Szekely, A., Jacobsen, T., D'Amico, S., Devescovi, A., Andonova, E., Herron, D.,
et al. (2004). A new on-line resource for psycholinguistic studies. *Journal
of Memory and Language*, 51(2), 247–250. doi:10.1016/j.jml.2004.03.002
- Tanaka, K. (1996). Inferotemporal cortex and object vision. *Neuroscience*, 19,
109–139. doi:10.1146/annurev.ne.19.030196.000545
- Thierry, G., & Wu, Y. J. (2007). Brain potentials reveal unconscious translation
during foreign-language comprehension. *Proceedings of the National
Academy of Sciences*, 104(30), 12530.
- 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(11), 4567–4570. doi:10.1073/pnas.0811155106
- Thierry, G., Cardebat, D., & Démonet, J.-F. (2003). Electrophysiological
comparison of grammatical processing and semantic processing of single
spoken nouns. *Cognitive Brain Research*, 17(3), 535–547. doi:10.1016/
S0926-6410(03)00168-X
- Thierry, G., Martin, C. D., Downing, P., & Pegna, A. J. (2007a). Controlling for
interstimulus perceptual variance abolishes N170 face selectivity. *Nature
Neuroscience*, 10(4), 505–511. doi:10.1038/nn1864
- Thierry, G., Martin, C., Downing, P., & Pegna, A. (2007b). Is the N170 sensitive
to the human face or to several intertwined perceptual and conceptual
factors? *Nature Neuroscience*, 10(7), 802–803. doi:10.1038/nn0707-802

- Turatto, M., Angrilli, A., Mazza, V., Umiltà, C., & Driver, J. (2002). Looking without seeing the background change: Electrophysiological correlates of change detection versus change blindness. *Cognition*, *84*(1), B1–B10.
- Uchikawa, K. K., & Boynton, R. M. R. (1987). Categorical color perception of Japanese observers: comparison with that of Americans. *Vision Research*, *27*(10), 1825–1833. doi:10.1016/0042-6989(87)90111-8
- van de Meerendonk, N., Kolk, H. H. J., Vissers, C. T. W. M., & Chwilla, D. J. (2010). Monitoring in language perception: mild and strong conflicts elicit different ERP patterns. *Journal of cognitive neuroscience*, *22*(1), 67–82. doi:10.1162/jocn.2008.21170
- Vigliocco, G., Vinson, D. P., Paganelli, F., & Dworzynski, K. (2005). Grammatical Gender Effects on Cognition: Implications for Language Learning and Language Use. *Journal of Experimental Psychology: General*, *134*(4), 501–520. doi:10.1037/0096-3445.134.4.501
- Vos, S. H. S., Gunter, T. C. T., Kolk, H. H. H., & Mulder, G. G. (2001). Working memory constraints on syntactic processing: an electrophysiological investigation. *Psychophysiology*, *38*(1), 41–63. doi:10.1111/1469-8986.3810041
- Waxman, S. (2001). Seeing Pink Elephants: Fourteen-Month-Olds' Interpretations of Novel Nouns and Adjectives. *Cognitive Psychology*, *43*(3), 217–242. doi:10.1006/cogp.2001.0764

- Waxman, S. R. S. (1999). Specifying the scope of 13-month-olds' expectations for novel words. *Cognition*, 70(3), B35–B50. doi:10.1016/S0010-0277(99)00017-7
- Waxman, S., & Booth, A. (2003). The origins and evolution of links between word learning and conceptual organization: new evidence from 11-month-olds. *Developmental Science*, 6(2), 128–135.
- Whorf, B. L., (1956). *Language, thought and reality*. Cambridge:MIT Press.
- Winawer, J. J., Witthoft, N. N., Frank, M. C. M., Wu, L. L., Wade, A. R. A., & Boroditsky, L. L. (2007). Russian blues reveal effects of language on color discrimination. *Proceedings of the National Academy of Sciences*, 104(19), 7780–7785. doi:10.1073/pnas.0701644104
- Winkler, I., & Czigler, I. (2012). Evidence from auditory and visual event-related potential (ERP) studies of deviance detection (MMN and vMMN) linking predictive coding theories and perceptual object representations. *International Journal of Psychophysiology*, 83(2), 132–143. doi:10.1016/j.ijpsycho.2011.10.001
- Winkler, I., Czigler, I., Sussman, E., Horváth, J., & Balázs, L. (2005). Preattentive binding of auditory and visual stimulus features. *Journal of cognitive neuroscience*, 17(2), 320–339.
- Wolff, P., & Holmes, K. J. (2011). Linguistic relativity. *Wiley Interdisciplinary Reviews: Cognitive Science*, 2(3), 253–265. doi:10.1002/wcs.104

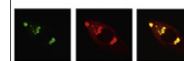
- Wu, Y. J., & Thierry, G. (2010). Chinese-English Bilinguals Reading English Hear Chinese. *Journal of Neuroscience*, 30(22), 7646–7651. doi:10.1523/JNEUROSCI.1602-10.2010
- Xu, F. (2002). The role of language in acquiring object kind concepts in infancy. *Cognition*, 85(3), 223–250. doi:10.1016/S0010-0277(02)00109-9
- Yee, E., Ahmed, S. Z., & Thompson-Schill, S. L. (2012). Colorless Green Ideas (Can) Prime Furiously. *Psychological Science*. doi:10.1177/0956797611430691
- Younger, B. A. B. (1985). The segregation of items into categories by ten-month-old infants. *Child Development*, 56(6), 1574–1583. doi:10.2307/1130476
- Zhang, S., & Schmitt, B. (1998). Language-dependent classification: The mental representation of classifiers in cognition, memory, and ad evaluations. *Journal of Experimental Psychology: Applied*, 4(4), 375.

Publications

1. Boutonnet, B., Athanasopoulos, P., & Thierry, G. (2012)

Available online at www.sciencedirect.com
SciVerse ScienceDirect
www.elsevier.com/locate/brainres

Brain Research



Research Report

Unconscious effects of grammatical gender during object categorisation

 Bastien Boutonnet^{a,c}, Panos Athanasopoulos^b, Guillaume Thierry^{a,c,*}
^aESRC Bilingualism Centre, Bangor, UK

^bSchool of Education, Communication & Language Sciences, Newcastle University, UK

^cSchool of Psychology, Bangor University, Bangor, UK

ARTICLE INFO

Article history:

Accepted 25 August 2012

Available online 31 August 2012

Keywords:

Linguistic relativity

Syntactic gender

Event-related brain potentials

Left anterior negativity

Object categorisation

ABSTRACT

Does language modulate perception and categorisation of everyday objects? Here, we approach this question from the perspective of grammatical gender in bilinguals. We tested Spanish–English bilinguals and control native speakers of English in a semantic categorisation task on triplets of pictures in an all-in-English context while measuring event-related brain potentials (ERPs). Participants were asked to press a button when the third picture of a triplet belonged to the same semantic category as the first two, and another button when it belonged to a different category. Unbeknownst to them, in half of the trials, the gender of the third picture name in Spanish had the same gender as that of the first two, and the opposite gender in the other half. We found no priming in behavioural results of either semantic relatedness or gender consistency. In contrast, ERPs revealed not only the expected semantic priming effect in both groups, but also a negative modulation by gender inconsistency in Spanish–English bilinguals, exclusively. These results provide evidence for spontaneous and unconscious access to grammatical gender in participants functioning in a context requiring no access to such information, thereby providing support for linguistic relativity effects in the grammatical domain.

© 2012 Elsevier B.V. All rights reserved.

1. Introduction

Whorf (1956), one of the two fathers of linguistic relativity, famously suggested that the language(s) one speaks shapes the way one thinks. The questions underpinning the linguistic relativity debate are questions such as: does language modulate perception? Is language encapsulated or does it interact with other cognitive processes? If so, what is the nature of these interactions and what properties of language bring these interactions to bear?

Scholars have misinterpreted Whorf's thesis as a formulation close to linguistic determinism, a far stronger claim that language may cause changes in basic physiological processes, e.g., that of visual perception. Pinker (1995), for instance, stated that “no matter how influential language may be, it would seem preposterous to a physiologist that it could reach down into the retina and rewire the ganglion cells (p. 63).” Over the past two decades, however, the linguistic relativity hypothesis has been resurrected in a milder form, perhaps closer to the original thinking of Sapir

*Correspondence to: School of Psychology, Brigantia Building, Bangor University, Bangor (Gwynedd) LL57 2AS, UK.

Fax: +44 1248 382 599.

E-mail address: g.thierry@bangor.ac.uk (G. Thierry).

Table 1 – Example of experimental conditions.

Picture primes		Targets	Gender congruency	Semantically relatedness
Tomato	Celery	Asparagus	+	+
Tomato	Celery	Carrot	–	+
Tomato	Celery	Truck	+	–
Tomato	Celery	Bike	–	–

and Whorf (Gumperz and Levinson, 1996; Gentner and Goldin-Meadow, 2003). Additionally, interactive models developed by McClelland and Rumelhart (1981), and Humphrey et al. (1997) and a recent, more direct application of these models to linguistic relativity by Lupyan (2012), offer working hypotheses as regards the cognitive mechanisms underpinning language-thought interactions. The label-feedback hypothesis, in particular, proposes that language is highly interconnected with other cognitive processes such as vision and categorisation and influences other functional networks in a top-down fashion.¹

Several recent studies have highlighted areas where lexical and grammatical information may affect cognitive processes other than language. Lexical characteristics of languages have been shown to affect colour perception in behavioural (Ozgen, 2004; Roberson et al., 2005; Franklin et al., 2008; Athanasopoulos, 2009) and neurophysiological (Thierry et al., 2009; Clifford et al., 2010; Liu et al., 2010) investigations. A number of studies have also reported effects of language in the domain of spatial representation and event conceptualisation where speakers exhibit differences in event description and recollection (Bowermann and Choi, 1991; Majid et al., 2004; Papafragou and Selimis, 2010) or even show different gaze patterns when exploring videos depicting events (Flecken, 2010). Studies investigating differences in grammatical number expression (e.g. languages with classifiers systems) have also suggested alteration of object classification (Lucy, 1992; Zhang and Schmitt, 1998; Athanasopoulos, 2007; Saalbach and Imai, 2007).

One interesting feature of some languages, which offers an appropriate test case for linguistic relativity, is grammatical gender. This feature, present in many of the world's languages, forces all nouns to be assigned to, most commonly, two or three classes: *masculine* and *feminine* and/or *neuter* (Corbett, 1991). Grammatical gender is of particular interest for two reasons: (a) when it is absent, it cannot be replaced by other lexicalisation patterns (unlike classifiers, for instance), and (b) its assignment is arbitrary except in the case of natural gender (male/female distinction).

With regard to point (a), for instance, Chinese requires the use of a marker before every quantified noun as in 'yi zhang zhi piao' (a [FLAT OBJECT] bank note), and English can sometimes do the same as in 'a piece of paper' or 'a flock of sheep'. By contrast, grammatical gender, when absent from a language, cannot be replaced by any combination of words.

¹ The label-feedback hypothesis is very recent and has only thoroughly been implemented in the context of label effects on object perception and categorisation. However, there is no reason to believe that the feedback processes it outlines do not arise for grammatical features such as gender.

Regarding point (b), the French word for 'sentry' (*une sentinelle*), for instance, is feminine, but rare must have been women sentries; a *toaster* is masculine in French (*un grille-pain*) but feminine in Spanish (*una tostadora*); in German a woman is feminine (*die Frau*) but a girl is neuter (*das Mädchen*). Even diachronically, the gender of nouns can change: the old word for girl in Polish used to be feminine (*ta dziewczyna*) but nowadays it is neuter (*to dziewczę*). In other words, both within a language and cross-linguistically, the relation between grammatical gender and word meaning appears to escape logic.

Studies investigating grammatical gender to date, have essentially focused on potential links between grammatical gender and object categorisation using (a) the voice-attribution paradigm (Sera et al., 1994; 2002; Bassetti, 2007; Forbes et al., 2008), (b) common noun–proper noun associations (Boroditsky et al., 2003; Phillips and Boroditsky, 2003), (c) semantic ratings and adjective associations (Flaherty, 2001; Boroditsky et al., 2003), or (d) a combination of the above methods (Vigliocco et al., 2005; Ramos and Roberson, 2010). Unfortunately, in all of these cases, the interpretation falls short of establishing the source of effects at an abstract level, disconnected from language itself. As Pinker (2007) put it: "speakers of different languages tilt in different directions in a woolly task, rather than having differently structured minds" (p. 148). The most recent and perhaps strongest evidence of grammatical gender-driven relativity comes from a study by Cubelli et al. (2011), who minimised the possibility that participants could use language as a strategy by using a non verbal semantic task on pictures. However, not all studies investigating implicit effects of grammatical gender on object categorisation have reported overwhelming evidence for such effects (see Kousta et al., 2008) and several have led to mixed results (Sera et al., 2002; Bassetti, 2007).

The greatest limitation of studies conducted so far in this field is their reliance on behavioural measurements. Indeed, as vigorously argued by Pinker (2007), behavioural evidence is open to contamination by explicit and/or idiosyncratic strategies used by participants to resolve the tasks, a process that is likely to solicit language processing (e.g., inner speech, sub-vocal rehearsal of instructions, covert denomination of objects, lexical access, etc.). If language access is prompted by the task at hand, then nothing can be said of the spontaneity of this effect. What is needed then is a method, which detects spontaneous access to grammatical gender representations without explicit involvement of language and not merely inferred from behavioural observations (Cf. Pinker, 2007).

In the present study, we asked participants to decide whether the third of a series of three objects presented one-by-one on the screen belonged to the same semantic

category as the two first ones. Semantic relatedness amongst the three objects was manipulated along with a covert manipulation of grammatical gender consistency (Table 1). We predicted that semantic incongruence would result in a modulation of the N400 wave, a negative-going potential with an average peak latency of 400 ms post-stimulus and known to reflect semantic integration mechanisms (Kutas and Hillyard, 1980; 1984). On the other hand, and critically, we hypothesised that grammatical gender inconsistency may modulate the Left-Anterior Negativity (LAN), a ERP marker of morphosyntactic processing (Friederici et al., 1993; Friederici and Jacobsen, 1999; Hahne and Friederici, 1999; Thierry, 2003). If such results were obtained, it would mean that grammatical gender is retrieved automatically and unconsciously rather than strategically and consciously during object categorisation.

2. Results

Behavioural data: Accuracy in the semantic relatedness task was overall high (79%) and was not studied by group or condition for the reason stated in Section 5 Fig. 1.

Regarding reaction times, we found a significant main effect of group ($F(1, 30)=11.4, p<0.002, \eta_p^2=0.28$) but we did not find any significant effect of semantic relatedness ($F(1, 30)=0.824, p>0.1$) or gender consistency ($F(1, 30)=0.010, p>0.1$) and no significant interactions between factors.

Electrophysiological data: N1 and P1 were unaffected by experimental conditions in either of the participant groups. As expected, the N4 was maximal over the centroparietal electrode sites and peaked at 361 ms on average (Fig. 2). There was a significant main effect of semantic relatedness on N400 mean amplitude between 300 and 400 ms ($F(1, 30)=40.74, p<0.0001, \eta_p^2=0.576$) such that the N400 was less negative in the semantically related than unrelated conditions (Bonferroni, $p<0.001$). There was no interaction between groups and the other experimental factors (Fig. 2A and B)

The LAN was maximal over left frontal regions and peaked at 559 ms on average (Fig. 3). There was a significant main effect of gender consistency ($F(1, 30)=23.5, p<0.0001,$

$\eta_p^2=0.439$) such that LAN amplitudes were more negative in the gender inconsistent than gender consistent condition (Bonferroni, $p<0.001$). This main effect was qualified by a significant interaction between gender and group ($F(1, 30)=4.97, p<0.05, \eta_p^2=0.142$). Post-hoc analyses revealed that the effect of gender was present for the Spanish–English bilinguals ($F(1, 14)=29.5, p<0.001, \text{Fig. 3A}$) but not for the English monolinguals ($F(1, 16)=3.1, p>0.05, \text{Fig. 3B}$).

Fig. 4 plots the p -value (negative log of 10, for presentation purposes) of the t -tests carried out millisecond by millisecond on the difference between gender consistent and gender inconsistent conditions. The difference between conditions became significant 388 ms after stimulus onset in the Spanish–English bilinguals and remained so until the end of the analysed epoch (700 ms) and was never significant for more than 30 ms in the English monolinguals (Rugg et al., 1995).

3. Discussion

The aim of the present study was to determine whether some features of language affect other cognitive processes such as object categorisation. Grammatical gender is a feature of some languages that has received considerable amounts of attention in linguistics and psycholinguistics but to our knowledge its effect on object categorisation has never been established based on measures of brain activity.

Unlike Cubelli et al. (2011) who used a similar experimental design, we found no behavioural effect of semantic relatedness of grammatical gender in Spanish–English bilingual participants. However, Cubelli et al. (2011) (a) used pairs rather than triads of pictures and (b) tested their participants in an Italian-speaking or Spanish-speaking environment (Italian and Spanish students tested at the university of Padova and Granada, respectively):

- (a) We used three pictures instead of two in order to (1) load participants' working memory, thereby increasing task difficulty and therefore reducing the likelihood of participants having enough executive resources to work out the

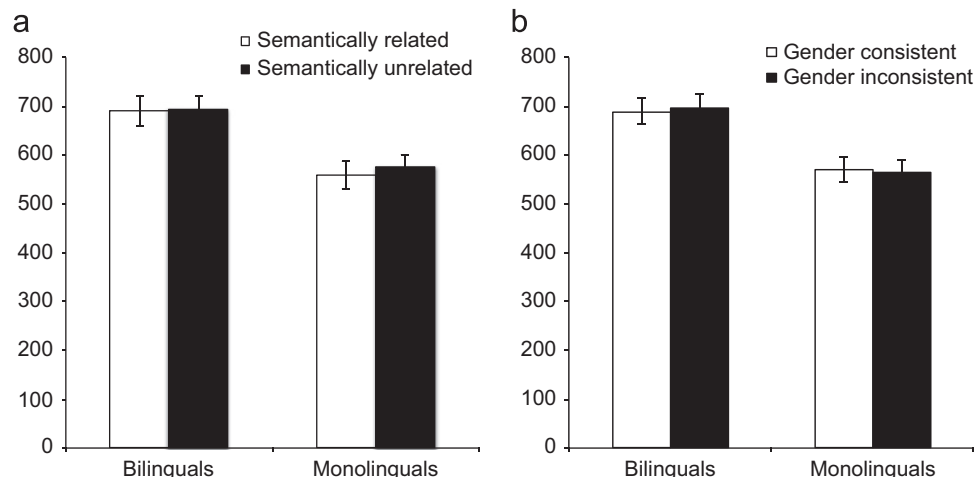


Fig. 1 – Plot of reaction times on correct trials only for the Spanish–English bilinguals and English monolinguals, showing no effect of condition. (A) Semantic conditions, (B) gender conditions. Error-bars depict s.e.m.

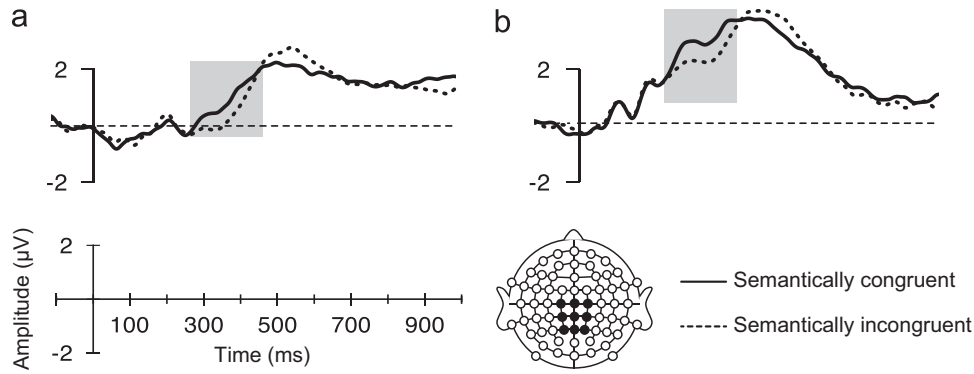


Fig. 2 – ERPs elicited in the semantic related and semantic unrelated conditions. (A) Spanish-English bilinguals, (B) English monolinguals. Linear derivation of electrodes C1, CZ, C2, CP1, CPZ, CP2, P1, PZ, P2.

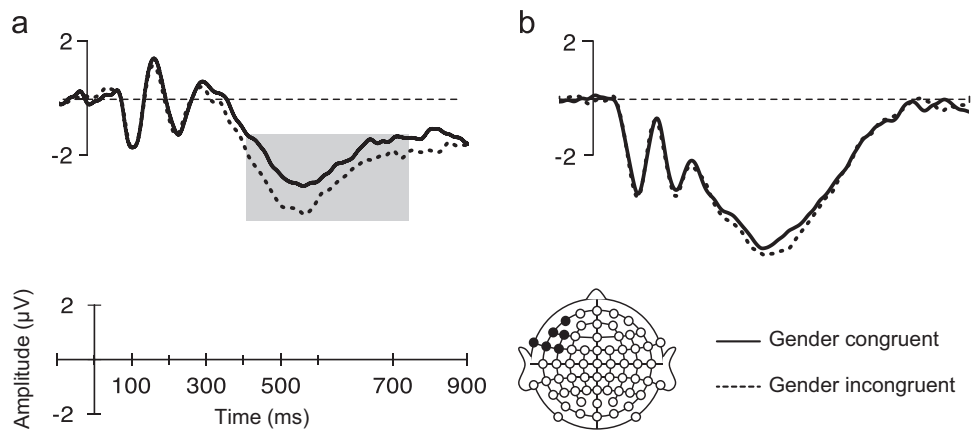


Fig. 3 – ERPs elicited in the gender consistent and gender inconsistent conditions. (A) Spanish-English bilinguals, (B) English monolinguals. Linear derivation of electrodes FT9, FT7, FC5, F7, F5, AF7.

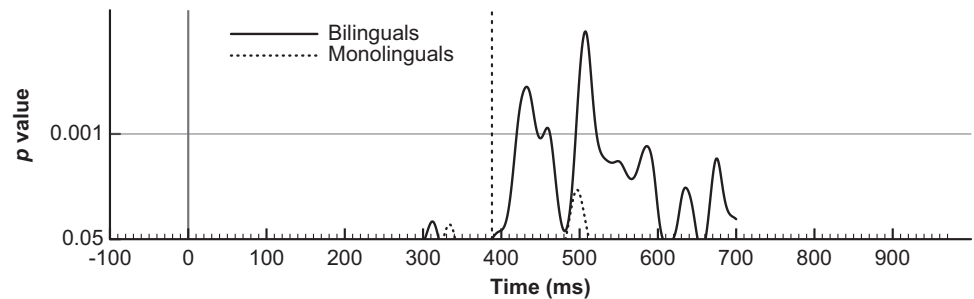


Fig. 4 – Plot of the p-value ms by ms for the LAN (negative log of 10). Dashed vertical line at 388 ms indicates when the effect becomes reliably significant.

hidden manipulation and (2) maximise the explicit and implicit priming effects (i.e., experimental sensitivity) due to the more consistent baseline produced by the first two pictures;

(b) The speaking environment has been shown to have a tangible impact on the language mode of individuals (Grosjean, 1998; Elston-Guttler et al., 2005). Our participants were tested in an all-in-English context during and outside the experimental session. It must be noted that Welsh is also spoken in the region of North-Wales but that exposure to Welsh is rare to very rare if individuals are not

actively seeking it, the medium of conversation being essentially English.

Nevertheless, we found the predicted semantic priming effect on N400 ERP amplitude (Kutas and Hillyard, 1980, 1984), showing that semantic priming amongst the picture triads was present even though it did not manifest itself behaviourally.

Critically, in addition to the semantic relatedness effect, we found a grammatical gender consistency effect in the ERP

data exclusively in the Spanish–English bilinguals, manifesting itself as a LAN modulation and showing that these participants extracted gender information while engaged in a semantic categorisation task requiring no such information. Since participants were tested in an all-in-English context, were never made aware of the gender manipulation, never reported being aware of it after debriefing, and since gender consistency had no behavioural effect, we interpret this result as evidence that access to gender information was implicit and unconscious (see [Thierry and Wu, 2007](#); [Wu and Thierry, 2010](#)). This result indicates that grammatical gender is spontaneously retrieved during semantic processing of pictures even though lexical-semantic processing was not explicitly required ([Strijkers et al., 2011](#)).

This result could be interpreted in terms of mere spreading of activation leading to the activation of grammatical gender representation even though accessing gender was irrelevant. Similarly, grammatical gender has been shown to affect picture naming cross-linguistically in bilingual word production ([Lemhöfer et al., 2008](#)) and there is good evidence that it is transferred in bilingualism ([Ganushchak et al., 2011](#)). However, we note that most studies having brought to light such spontaneous effects of grammatical gender retrieval have used tasks that rely heavily on language activation. This was not the case in the current study since participants were not required to name pictures or retrieve any verbal information to perform semantic categorisation. In addition, the absence of a behavioural effect in our study suggests that spreading activation alone is not sufficient to account for the pattern of results obtained.

Irrespective of the fact that we did not find any behavioural effects, our conclusions are similar to those of [Cubelli et al. \(2011\)](#) with the added dimension that such effect is probably encountered on an unconscious level. Altogether our results suggest that object conceptual retrieval and categorisation are unconsciously affected by language-specific syntactic information, such as grammatical gender, even when such information is task-irrelevant. Similar results of access to task-irrelevant semantic features were obtained by [Yee et al. \(2012\)](#). The demonstration of this phenomenon in the grammatical domain supports the view that language substantially interacts with other cognitive processes, and further highlights the critical role of language in shaping the way humans process reality and the world around them.

This conclusion is inconsistent with the modularity of language hypothesis ([Fodor, 1975](#); [Chomsky, 2000](#); [Fodor, 2008](#)) and rather suggests that the organisation of information at the cortical level relies heavily on interconnectivity and interactions amongst distributed cell assemblies

([McClelland and Rumelhart, 1981](#); [Humphreys et al., 1997](#); [Pulvermüller, 1999](#); see [Martin, 2007](#) for review). The data also lend support to the linguistic relativity hypothesis, and its newest development ([Lupyan, 2012](#)), by showing that semantic features of objects are spontaneously retrieved together with semantically irrelevant information such as syntactic gender and this information likely contributes to participants' mental representations of these objects.

4. Conclusion

While language does not necessarily determine thoughts, and while thinking may be possible without the aid of language, it nonetheless provides a ready basis of information for the purposes of classifying the world into meaningful categories ([Lucy, 1997](#)). To date, this observation has been empirically demonstrated primarily in the domain of colour ([Regier and Kay, 2009](#); [Thierry et al., 2009](#); [Clifford et al., 2010](#); [Liu et al., 2010](#)). The current study shows that humans may automatically utilise grammatical categories such as gender when asked to make judgements about semantic relationships unrelated to the grammatical categories in question. The fact that we have found such effects in the domain of grammatical gender is particularly important, since previous empirical attempts to address the Whorfian question in this domain used methods and task instructions, mostly based on behavioural measures, that might promote strategic use of grammatical gender categories. Future studies will shed more light on the locus of this effect as well as patterns of brain connectivity, and establish whether it generalises to other language-specific properties, e.g., compound words, classifiers, and highly grammaticised language features, such as tense and aspect.

5. Experimental procedures

Participants: Participants were 16 Spanish native speakers with English as a second language (L2) (henceforth Spanish–English bilinguals). One participant was eliminated due to insufficient data quality. The 15 remaining participants (9 female, age: 32.6 ± 1.981 ; 6 male, age: 29.3 ± 2.028) were included in the final analyses. All the participants learned English at least in primary and secondary school in Spanish speaking countries and were living in the UK at the time of testing. [Table 2](#) summarises participants' language experience and self-assessed proficiency in L1 and L2. At the time of testing, the participants were using L2 slightly more

Table 2 – Characteristics of Spanish–English participants.

Measure	Mean	Standard error
L1 self-rating (10 pt scale)	9.7	(0.1)
L2 self-rating (10 pt scale)	8.5	(0.4)
Daily L1 usage (%)	44.6	(6.5)
Daily L2 usage (%)	55.4	(6.5)
Age of L2 acquisition (years)	10.2	(1)
Length of immersion (months)	52.6	(12.3)

than L1, due to their immersion context but the difference in self-reported use was not significant. Proficiency was however significantly higher in L1 than L2 ($z = -2.49$, $p < 0.05$).

Twenty native speakers of English who all reported that they were monolingual also took part in the experiment as a control group. Three of them were eliminated due to insufficient data quality. The 17 remaining participants (9 female, age: 19.57 ± 0.53 ; 8 male, age: 19.25 ± 0.16) were included in the final analyses. All the participants were right-handed and had normal or corrected-to-normal vision.

Materials: We selected 288 black-and-white line drawings from Snodgrass and Vanderwart (1980) and Szekely et al. (2004). Pictures were grouped in 96 triads, such that the name of two first pictures had the same gender and they belonged to the same semantic category and the third picture was either from the same or a different semantic category and their name either had the same or the opposite gender as the two others. The 96 triads could therefore be split into 48 semantically related and 48 semantically unrelated associations, or 48 gender consistent and 48 gender inconsistent associations, providing 24 triads per individual experimental conditions.

Procedure: After participants filled out a questionnaire about their language learning background and self-assessed proficiency in L1 and L2, they were tested individually in a quiet room. They were seated in front of a computer monitor (CRT make 19", 100 cm from the screen) on which picture stimuli were displayed within a viewing angle of 8° and handed a response box. The participants were instructed to press a given button if the three pictures of a triad belonged to the same semantic category and another button if not. Participants were never told about the covert gender manipulation. On each trial, a fixation cross was presented for 1000 ms, followed immediately by the first prime for a duration of 600 ms, then the second picture appeared after a blank screen of 250 ms duration, for a duration 600 ms. Then the target (third picture) appeared after a variable interval randomly selected between 300 and 500 ms in steps of 50 ms, in order to cancel offset effects. The target remained on screen until participants responded. Five practice trials preceded the experimental trials. All experimental instructions were provided in English. The order of blocks was counterbalanced across participants and the presentation of items was randomised within each block.

Electrophysiological recording: The EEG was continuously recorded at a rate of 1 kHz from 64 Ag/AgCl electrodes placed according to the extended 10–20 convention. Two additional electrodes were attached above and below the left eye and on either side of the left and right eye in order to monitor for eye-blinks and horizontal eye movements. Cz was the reference electrode during acquisition. Impedances were maintained below 5 k Ω for all 64 electrodes and below 10 k Ω for vertical electrooculogram electrodes. EEG signals were filtered off-line using a 30 Hz low pass zero phase shift digital filter.

Behavioural data analysis: Given that the task performed by participants was a semantic relatedness task, accuracy was not informative regarding access to gender information, especially if we consider that some semantic associations were far-fetched due to the necessity of creating a fully counter-balanced experimental design. Indeed, some of the triads may have seemed unrelated to the participant when they were considered related by the experimenters

and vice versa. Therefore, we considered reaction times irrespective of response accuracy and we did not consider potential differences in accuracy arising between groups or conditions. However, RTs shorter than 250 ms and differing by more than 2.5 standard deviations from the average RT in each condition and participant were individually discarded. A two-way ANOVA by participant was conducted on the RTs with semantic relatedness (related, unrelated), grammatical gender consistency (consistent, inconsistent) as within-subject factors and group (Spanish–English bilinguals, English monolinguals) as between-subject factor.

Electrophysiological data analysis: Eye-blink artefacts were mathematically corrected using the algorithm provided in Scan 4.4™. The algorithm is derived from the method advocated by Gratton et al. (1983). Note that eye-blink occurred mostly after the response was made as a consequence of special instruction given to the participants. ERPs were then computed by averaging EEG epochs ranging from -100 to 1000 ms after stimuli onset. Baseline correction was applied in relation to 100 of pre-stimulus activity and individual averages were re-referenced to the global field power produced over the entire scalp. ERPs time-locked to the onset of target pictures were visually inspected and mean amplitudes were measured in temporal windows determined based on variations of the mean global field power measured across the scalp (Picton et al., 2000). Four components were identified as expected. The P1 and N1 were maximal at parietal sites and were measured in the 100–150 ms range for the P1 and 170–230 ms for the N2, the N400 was maximal on central sites (Cz) and was measured in the 300–400 ms window. Finally, the left anterior negativity, strongest at left anterior recording sites, was measured in the 380–600 ms window. Peak latencies were measured at sites of maximal amplitude (PO8 for the P1 and N1, CZ for the N4; FT9 for the LAN) and mean ERP amplitudes were measured in regions of interest around the sites of maximal amplitude (O1, PO3, PO7, O2, PO4, PO8 for the P1 and N1; C1, CZ, C2, CP1, CPZ, CP2, P1, PZ, P2 for the N4; FT9, FT7, FC5, F7, F5, AF7 for the LAN). Note that we did not conduct a full-scalp analysis because the modulation of the ERP components were predicted to occur in the regions of interest and therefore statistical analyses of ERP mean amplitude were conducted in sets of electrodes determined *a priori* based on the LAN and N400 literature (cf. introduction). Mean amplitudes and peak latencies were subjected to a mixed repeated-measures ANOVA with semantic relatedness (related, unrelated), grammatical gender consistency (consistent, inconsistent) and electrode (6 or 9 levels) as within-subject factors and group (Spanish–English bilinguals, English monolinguals) as between-subject factor. In addition, paired sample t-tests were conducted between the gender consistent and gender inconsistent conditions millisecond-by-millisecond to determine the onset of differences between conditions (using a linear derivation of the 6 electrodes used in the mean amplitude analysis).

Acknowledgments

B.B. is funded by the Economic and Social Research Council (RES-592-28-0001). G.T. is funded by the Economic and Social Research Council (RES-000-23-0095) and the European Research Council (ERC-StG-209704).

REFERENCES

- Athanasopoulos, P., 2007. Interaction between grammatical categories and cognition in bilinguals: the role of proficiency, cultural immersion, and language of instruction. *Lang. Cogn. Process.* 22, 689–699.
- Athanasopoulos, P., 2009. Cognitive representation of colour in bilinguals: the case of Greek blues. *Bilingualism* 12, 83.
- Bassetti, B., 2007. Bilingualism and thought: grammatical gender and concepts of objects in Italian–German bilingual children. *Int. J. Bilingualism* 11, 251–273.
- Boroditsky, L., Schmidt, L.A., Phillips, W., 2003. Sex, syntax, and semantics. *Language in Mind: Advances in the Study of Language and Cognition*, pp. 61–79.
- Bowermann, M., Choi, S., 1991. Learning to express motion events in English and Korean: the influence of language-specific lexicalization patterns. *Cognition* 41, 83–121.
- Chomsky, N., 2000. *New Horizons in the Study of Language and Mind*. Cambridge University Press, Cambridge [England], New York.
- Clifford, A., Holmes, A., IRL, Davies, Franklin, A., 2010. Colour categories affect pre-attentive colour perception. *Biol. Psychol.* 85, 275–282.
- Corbett, G., 1991. *Gender*. Cambridge University Press.
- Cubelli, R., Paolieri, D., Lotto, L., Job, R., 2011. The effect of grammatical gender on object categorisation. *J. Exp. Psychol. Learn. Mem. Cogn.* 37, 449–460.
- Elston-Guttler, K., Gunter, T., Kotz, S., 2005. Zooming into L2: global language context and adjustment affect processing of interlingual homographs in sentences. *Brain Res. Cogn. Brain Res.* 25, 57–70.
- Flaherty, M., 2001. How a language gender system creeps into perception. *J. Cross-Cult. Psychol.* 32, 18–31.
- Flecken, M., 2010. Event conceptualization by early Dutch–German bilinguals: insights from linguistic and eye-tracking data. *Bilingualism* 14, 61–77.
- Fodor, J.A., 1975. *The Language of Thought*. Harvard University Press, Cambridge, Mass.
- Fodor, J.A., 2008. *LOT 2: The Language of Thought Revisited*. Clarendon Press; Oxford University Press, Oxford; New York.
- Forbes, J.N., Poulin-Dubois, D., Rivero, M.R., Sera, M.D., 2008. Grammatical gender affects bilinguals' conceptual gender: implications for linguistic relativity and decision making. *Open Appl. Linguist. J.* 1, 68–76.
- Franklin, A., Drivonikou, G., Clifford, A., Kay, P., Regier, T., Davies, I., 2008. Lateralization of categorical perception of colour changes with colour term acquisition. *Proc. Natl. Acad. Sci. USA* 105, 18221.
- Friederici, A.D., Jacobsen, T., 1999. Processing grammatical gender during language comprehension. *J. Psycholinguist. Res.* 28, 467–484.
- Friederici, A.D., Pfeifer, E., Hahne, A., 1993. Event-related brain potentials during natural speech processing: effects of semantic, morphological and syntactic violations. *Brain Res. Cogn. Brain Res.* 1, 183–192.
- Ganushchak, L.Y., Verdonschot, R.G., Schiller, N.O., 2011. When leaf becomes neuter. *Neuroreport*, event-related potential evidence for grammatical gender transfer in bilingualism. *Neuroreport* 22, 106–110.
- Gentner, D., Goldin-Meadow, S., 2003. *Language in Mind: Advances in the Study of Language and Thought*. MIT Press, Cambridge, Mass.
- Gratton, G., Coles, M.G.H., Donchin, E., 1983. A new method for off-line removal of ocular artifact. *Electroencephalogr. Clin. Neurophysiol.* 55, 468–484.
- Grosjean, F., 1998. Transfer and language mode. *Bilingualism* 1, 175–176.
- Gumperz, J.J., Levinson, S.C., 1996. *Rethinking Linguistic Relativity*. Cambridge University Press, Cambridge; New York, NY, USA.
- Hahne, A., Friederici, A.D., 1999. Electrophysiological evidence for two steps in syntactic analysis. *Early automatic and late controlled processes*. *J. Cogn. Neurosci.* 11, 194–205.
- Humphreys, G.W., Riddoch, M.J., Price, C.J., 1997. Top-down processes in object identification: evidence from experimental psychology, neuropsychology and functional anatomy. *Philos. Trans. R. Soc. London B Biol. Sci.* 352, 1275–1282.
- Kousta, S.-T., Vinson, D.P., Vigliocco, G., 2008. Investigating linguistic relativity through bilingualism: the case of grammatical gender. *J. Exp. Psychol. Learn. Mem. Cogn.* 34, 843–858.
- Kutas, M., Hillyard, S.A., 1980. Reading senseless sentences: brain potentials reflect semantic incongruity. *Science* 207, 203–205.
- Kutas, M., Hillyard, S.A., 1984. Brain potentials during reading reflect word expectancy and semantic association. *Nature*, 161–163.
- Lemhöfer, K., Spalek, K., Schriefers, H., 2008. Cross-language effects of grammatical gender in bilingual word recognition and production. *J. Mem. Lang.* 59, 312–330.
- Liu, Q., Li, H., Campos, J.L., Teeter, C., Tao, W., Zhang, Q., Sun, H.-J., 2010. Language suppression effects on the categorical perception of colour as evidenced through ERPs. *Biol. Psychol.* 85, 45–52.
- Lucy, J.A., 1992. *Grammatical Categories and Cognition: A Case Study of the Linguistic Relativity Hypothesis*. Cambridge University Press, Cambridge.
- Lucy, J.A., 1997. Linguistic relativity. *Annu. Rev. Anthropol.* 26, 291–312.
- Lupyan, G., 2012. Linguistically modulated perception and cognition: the label-feedback hypothesis. *Front. Psychol.*, 1–13.
- Majid, A., Bowerman, M., Kita, S., Haun, D.B.M., Levinson, S.C., 2004. Can language restructure cognition? The case for space. *Trends Cogn. Sci.* 8, 108–114.
- Martin, A., 2007. The representation of object concepts in the brain. *Annu. Rev. Psychol.* 58, 25–45.
- McClelland, J.L., Rumelhart, D.E., 1981. An interactive activation model of context effects in letter perception: I. An account of basic findings. *Psychol. Rev.* 88, 375.
- Ozgen, E., 2004. Language, learning, and colour perception. *Curr. Dir. Psychol. Sci.* 13, 95–98.
- Papafraçou, A., Selimis, S., 2010. Event categorisation and language: a cross-linguistic study of motion. *Lang. Cogn. Process.* 25, 224–260.
- Phillips, W., Boroditsky, L., 2003. Can quirks of grammar affect the way you think? Grammatical gender and object concepts. In: *Proceedings of the 25th Annual Meeting of the Cognitive Science Society*, pp. 928–933.
- Picton, T., Bentin, S., Berg, P., Donchin, E., Hillyard, S., Johnson Jr., R., Miller, G., Ritter, W., Ruchkin, D., Rugg, M., 2000. Guidelines for using human event-related potentials to study cognition: recording standards and publication criteria. *Psychophysiology* 37, 127–152.
- Pinker, S., 1995. *The Language Instinct: The New Science of Language and Mind*. Penguin, London.
- Pinker, S., 2007. *The Stuff of Thought: Language as a Window into Human Nature*. Viking, New York.
- Pulvermüller, F., 1999. Words in the brain's language. *Behav. Brain Sci.* 22, 253–279.
- Ramos, S., Roberson, D., 2010. What constrains grammatical gender effects on semantic judgements? Evidence from Portuguese. *Eur. J. Cogn. Psychol.*, 1–16.
- Regier, T., Kay, P., 2009. Language, thought, and colour: Whorf was half right. *Trends Cogn. Sci.* 13, 439–446.
- Roberson, D., Davidoff, J., Davies, I.R.L., Shapiro, L.R., 2005. Colour categories: evidence for the cultural relativity hypothesis. *Cogn. Psychol.* 50, 378–411.

- Rugg, M.D., Doyle, M.C., Wells, T., 1995. Word and nonword repetition within-and across-modality: an event-related potential study. *J. Cogn. Neurosci.* 7, 209–227.
- Saalbach, H., Imai, M., 2007. Scope of linguistic influence: does a classifier system alter object concepts?. *J. Exp. Psychol. Gen.* 136, 485–501.
- Sera, M.D., Berge, C.A.H., JDC, Pintado, 1994. Grammatical and conceptual forces in the attribution of gender by English and Spanish speakers. *Cogn. Dev.* 9, 261–292.
- Sera, M.D., Elieff, C., Forbes, J., Burch, M.C., Rodríguez, W., Dubois, D.P., 2002. When language affects cognition and when it does not: an analysis of grammatical gender and classification. *J. Exp. Psychol. Gen.* 131, 377–397.
- Snodgrass, J.G., Vanderwart, M., 1980. A standardized set of 260 pictures: norms for name agreement, image agreement, familiarity, and visual complexity. *J. Exp. Psychol. Hum. Learn.* 6, 174.
- Strijkers, K., Holcomb, P.J., Costa, A., 2011. Conscious intention to speak proactively facilitates lexical access during overt object naming. *J. Mem. Lang.* 65, 345–362.
- Szekely, A., Jacobsen, T., D'Amico, S., Devescovi, A., Andonova, E., Herron, D., Lu, C.C., Pechmann, T., Pléh, C., Wicha, N., Federmeier, K., Gerdjikova, I., Gutierrez, G., Hung, D., Hsu, J., Iyer, G., Kohnert, K., Mehotcheva, T., Orozco-Figueroa, A., Tzeng, A., Tzeng, O., Arévalo, A., Vargha, A., Butler, A.C., Buffington, E., Bates, E., 2004. A new on-line resource for psycholinguistic studies. *J. Mem. Lang.* 51, 247–250.
- Thierry, G., 2003. Electrophysiological comparison of grammatical processing and semantic processing of single spoken nouns. *Cogn. Brain Res.* 17, 535–547.
- Thierry, G., Athanasopoulos, P., Wiggett, A., Dering, B., Kuipers, J.R., 2009. Unconscious effects of language-specific terminology on preattentive colour perception. *Proc. Natl. Acad. Sci. USA* 106, 4567–4570.
- Thierry, G., Wu, Y.J., 2007. Brain potentials reveal unconscious translation during foreign-language comprehension. *Proc. Natl. Acad. Sci. USA* 104, 12530.
- Vigliocco, G., Vinson, D.P., Paganelli, F., Dworzynski, K., 2005. Grammatical gender effects on cognition: implications for language learning and language use. *J. Exp. Psychol. Gen.* 134, 501–520.
- Whorf, B.L., Carroll, J.B., Chase, S., 1956. *Language, Thought and Reality*. Wiley, New York.
- Wu, Y.J., Thierry, G., 2010. Chinese–English bilinguals reading English hear Chinese. *J. Neurosci.* 30, 7646–7651.
- Yee, E., Ahmed, S.Z., Thompson-Schill, S.L., 2012. Colourless green ideas (can) prime furiously. *Psychol. Sci.*
- Zhang, S., Schmitt, B., 1998. Language-dependent classification: the mental representation of classifiers in cognition, memory, and ad evaluations. *J. Exp. Psychol. Appl.* 4, 375.

2. Boutonnet, B, Dering, B., Nuñas-Guasch, N., & Thierry, G. (2013)

Seeing Objects through the Language Glass

Bastien Boutonnet¹, Benjamin Dering², Nestor Viñas-Guash¹,
and Guillaume Thierry¹

Abstract

Recent streams of research support the Whorfian hypothesis according to which language affects one's perception of the world. However, studies of object categorization in different languages have heavily relied on behavioral measures that are fuzzy and inconsistent. Here, we provide the first electrophysiological evidence for unconscious effects of language terminology on object perception. Whereas English has two words for cup and mug, Spanish labels those two objects with the word "taza."

We tested native speakers of Spanish and English in an object detection task using a visual oddball paradigm, while measuring event-related brain potentials. The early deviant-related negativity elicited by deviant stimuli was greater in English than in Spanish participants. This effect, which relates to the existence of two labels in English versus one in Spanish, substantiates the neurophysiological evidence that language-specific terminology affects object categorization. ■

INTRODUCTION

The question of language–thought interactions has recently become a major topic of interest in cognitive neuroscience. It has become essential because of the debate on language encapsulation and on the potential effects of language on other cognitive processes (Fodor, 1975, 2008; Chomsky, 2000). The linguistic relativity hypothesis has undergone several interpretations since its inception by Whorf, Carroll, and Chase (1956). One early (misleading) interpretation of the hypothesis contends that language determines thought and, therefore, that without language thought is impossible. In light of compelling evidence that high-level cognitive operations are possible without language, this position has simply become untenable (e.g., number cognition in primates [Gallistel, 1989], infants [Feigenson, Dehaene, & Spelke, 2004], and in languages that do not have a complex lexicalized number system [Gordon, 2004]). On the other hand, recent theoretical reconceptualizations (e.g., Gentner & Goldin-Meadow, 2003; Gumperz & Levinson, 1996) have put forward nondeterministic versions of the hypothesis, according to which language influences (rather than determines) thought. The linguistic relativity debate has therefore moved toward the question of interaction between language representations and perception rather than that of determinism (Lucy, 1992a). However, this reconceptualization lacks psychological and physiological underpinning.

Here, we aimed at testing the validity of the most recent theoretical take on the Whorfian hypothesis, which does away with a "strong/weak" distinction (Klemfuss, Prinzmetal, & Ivry, 2012; Lupyan, 2012) and offers researchers clearer working hypotheses regarding language–thought interactions. For instance, based on interactive-processing models such as those developed by McClelland and Rumelhart (1981), the label–feedback hypothesis (Lupyan, 2012) proposes that language is highly interconnected with other cognitive processes such as vision and categorization and that it produces transient modulations of on-going perceptual (and higher level) processing. Whorfian effects can therefore arise from interactions among distributed brain regions, as in the case of prefrontal areas preparing the visual cortex to perceive particular dimensions of stimuli before they are actually displayed (Lamme & Roelfsema, 2000). This model therefore allows for nontrivial linguistic relativity effects to arise but is not tied in a deterministic view where perceptual areas are functionally structured by language (for an exhaustive explanation of the hypothesis and a review of the experimental literature, see Lupyan, 2012).

Previous studies have highlighted areas where lexical and grammatical information affect domain-general cognitive processes. For instance, lexicalization constraints on spatial representation and event conceptualization (e.g., focus on manner vs. end point of motion) have been shown to affect speakers' event description and recollection (Papafragou & Selimis, 2010; Majid, Bowerman, Kita, Haun, & Levinson, 2004; Bowerman & Choi, 1991) or to elicit different gaze patterns when exploring scenes depicting events (Flecken, 2010). Studies investigating grammatical number (i.e., language with classifier systems) reveal a tendency to categorize objects on the basis of

¹Bangor University, ²The Salk Institute for Biological Studies, La Jolla, CA

substance rather than shape when classifiers put the focus on substance (Saalbach & Imai, 2007; Zhang & Schmitt, 1998; Lucy, 1992b). In a similar vein, grammatical gender has also been shown to affect speakers' object categorization in covert gender assignment tasks (Kurinski & Sera, 2010; Forbes, Poulin-Dubois, Rivero, & Sera, 2008; Bassetti, 2007; Sera et al., 2002), judgment and adjective-association tasks (Boroditsky, Schmidt, & Phillips, 2003; Phillips & Boroditsky, 2003), and priming paradigms (Boutonnet, Athanasopoulos, & Thierry, 2012; Cubelli, Paolieri, Lotto, & Job, 2011). Finally, differences in color terminology have been shown to affect color perception in behavioral (Athanasopoulos, 2009; Franklin et al., 2008; Roberson, Davidoff, Davies, & Shapiro, 2005; Ozgen, 2004) and neurophysiological (Athanasopoulos, Dering, Wiggett, Kuipers, & Thierry, 2010; Clifford, Holmes, Davies, & Franklin, 2010; Liu et al., 2010; Thierry, Athanasopoulos, Wiggett, Dering, & Kuipers, 2009) investigations.

Despite the evidence in favor of the existence of Whorfian effects, it remains that studies in the field have mostly relied on behavioral measures. The problem is that such measures are open to contamination by explicit strategies used by participants to resolve the tasks, a process likely to involve language processing. Here, following neurophysiological investigations in the domain of color (Liu et al., 2010; Thierry et al., 2009; Franklin et al., 2008; Roberson, Pak, & Hanley, 2008; Gilbert, Regier, Kay, & Ivry, 2006), we investigated whether language-specific terminology also constrains object categorization (Gilbert, Regier, Kay, & Ivry, 2008). For instance, Thierry and colleagues (2009) recorded ERP correlates of color change detection in Greek–English bilinguals who have two color terms for blue (*ble* = “dark blue” and *ghalazio* = “light blue”) and a control group of English monolinguals. They found that native speakers of Greek exhibited a greater visual MMN (vMMN) elicited by blue deviants than English controls. On the basis of this paradigm, we chose to extend the evidence from the domain of color perception to that of object categorization. We chose drinking vessels because they have been examined thoroughly in previous cross-linguistic naming studies (Pavlenko & Malt, 2010; Ameel, Malt, Storms, & Van Assche, 2009; Ameel, Storms, Malt, & Sloman, 2005). These studies suggest that bilingual speakers' categorical boundaries shift through exposure to their second language—a phenomenon that has also been reported for colors (Athanasopoulos et al., 2010).

We recorded brain potentials from Spanish and English native speakers while they performed an object detection task within an oddball paradigm to test the extent to which unconscious aspects of visual object processing are modulated by one's language.

Spanish differs from English in the way some objects are labeled. Whereas English has two words to refer to a cup and a mug, Spanish only uses one label for these two objects: “taza.” In this experiment, participants were presented with three stimuli within an oddball paradigm (one of high local probability, i.e., standard; and two of

low local probability, i.e., deviants). Participants were instructed to detect a particular deviant stimulus or target (a bowl) in each of two experimental blocks. In one block, the nontarget deviant was a cup and the standard was a mug, and in the other block, the nontarget deviant was a mug and the standard was a cup.

We expected nontarget deviants to spontaneously elicit a deviant-related negativity (DRN) regardless of a response from the participants (Winkler, Czigler, Sussman, Horváth, & Balázs, 2005; Czigler, Balázs, & Pató, 2004; Czigler, Balázs, & Winkler, 2002; Turatto, Angrilli, Mazza, Umilta, & Driver, 2002; Csibra, Czigler, & Ambro, 1994). Because of the terminological difference between English and Spanish, we expected that the change from cup to mug would elicit a greater DRN in English than Spanish participants.

METHODS

Participants

Participants were 13 native speakers of Spanish (10 women, three men; $M_{Age} = 21$ years, $SD = 1.6$ years) tested in Spain and 14 native speakers of English tested in Wales (eight women, six men; $M_{Age} = 20$ years, $SD = 0.6$ years). Spanish participants were recruited from a database filtered to have a level no higher than A2 in English and a daily use of English lower than 5%. As part of the normal education curriculum in Spain, all Spanish participants received some exposure to English, but all reported having a limited knowledge of the language as well as a rare use of it. None of the Spanish participants had spent more than 2 weeks in an English-speaking country. The language usage background data used to filter the database were collected from self-reports from the participants before entry in the database.

Some of the Spanish speakers were also fluent in Catalan. This was not considered a problem because Catalan and Spanish are matched with respect to object denomination for cups and mugs. Some of the English participants reported having basic knowledge of other languages (including Spanish) but had self-reported very low proficiency and were not using any of their other languages on an everyday basis.

Materials

Three grayscale photographs of a cup, a mug, and a bowl subtending approximately 8° of visual angle were presented in the middle of a white background square in the center of a CRT monitor.

Procedure

Participants viewed two blocks of 450 stimuli. Within each block, a standard stimulus was presented with a high local

probability (either a cup or a mug, 80%). Deviant stimuli, presented with a low local probability, were either to be ignored (a mug or a cup, depending on the nature of the standard, 15%) or to be reported (bowl target, 5%). Presentation order was pseudorandomized such that two deviants or targets never appeared in immediate succession, and there were at least three standards in a row between two deviants. Stimuli were presented for 300 msec with a random variable ISI of 400, 450, 500, 550, and 600 msec, averaging to 500 msec. Participants were instructed to detect the target object (bowl) by pressing a button on a response box as quickly as possible. Block order was fully counterbalanced between participants.

Electrophysiological Recording

Electrophysiological data were recorded in two different laboratories. The Spanish participants were tested in Barcelona, Spain (Pompeu Fabra University). EEG was recorded (BrainVision Recorder 1.10) in reference to the left mastoid electrode at the rate of 1 kHz from 34 tin electrodes placed according to the 10–20 convention. Impedances were kept below 5 k Ω for electrodes on the cap and below 10 k Ω for external electrodes. The English participants were tested (NeuroScan 4.4) in Bangor, Wales (Bangor University). EEG was recorded in reference to the left mastoid electrode at the rate of 1 kHz from 34 Ag–Cl electrodes placed according to the 10–20 convention. All impedances were kept below 5 k Ω for electrodes on the cap and below 10 k Ω for external electrodes. Both data sets were analyzed using BrainVision Analyzer 2. EEG activity was filtered off-line with a low-pass 0.5-Hz filter (slope of 12 dB/oct) and a high-pass 30-Hz filter (slope of 48 dB/oct).

Data Analysis

Accuracy scores and RTs were submitted to independent samples *t* tests between groups (t_1 and t_2 , respectively). Eye blinks were mathematically corrected using the Gratton, Coles, and Donchin (1983) algorithm provided in Brain Vision Analyzer 2, and epochs with activity exceeding ± 75 μ V at any electrode site were automatically discarded. Epochs ranged from -100 to 600 msec after stimulus onset. Baseline correction was performed in reference to prestimulus activity, and individual averages were rereferenced to the left and right mastoid off-line. ERPs time-locked to the onset of the pictures were visually inspected, and mean amplitudes were measured in temporal windows determined based on variation of the mean global field power measured across the scalp (Picton et al., 2000). ERPs elicited by standard stimuli were averaged across blocks as were ERPs elicited by deviants; therefore, comparisons between standard and deviants did not reflect inherent perceptual differences between cups and mugs but only the deviancy effect.

Potential perceptual differences between the cup and mug objects were also investigated by analyzing amplitude and latency of the P1 peak from ERPs computed from standard stimuli, separately for each of the two experimental blocks. The P1 was maximal at parietal sites and was measured in the 100- to 150-msec range. Mean amplitude and latency of the P1 collected from a linear derivation of the five electrodes of interest (PO1, PO2, O1, OZ, and O2) were submitted to a 2 within-subject \times 2 between-subject ANOVA with standard object (cup/mug) as a within-subject factor and language group (Spanish/English) as a between-subject factor.

The DRN was defined as the earliest modulation of the negative component following the P1 over occipital recording sites. DRN analysis was conducted on individual ERPs elicited by standards and nontarget deviants, was maximal over the parieto-occipital scalp, and was studied in the 145- to 180-msec range at electrodes PO1, PO2, O1, OZ, and O2, predicted to be the electrodes of maximal sensitivity for the effect measured (Liu et al., 2010; Thierry et al., 2009). Mean amplitudes of ERPs from standard and deviant stimuli were subjected to a mixed repeated measures ANOVA with deviancy (deviant/standard) and electrode (five levels) as a within-subject factors and language group (Spanish/English) as a between-subject factor. In addition, paired sample *t* tests were conducted between the standard and deviant conditions millisecond-by-millisecond to determine the onset of differences between conditions (using a linear derivation of the five electrodes used in the mean amplitude analysis).

Furthermore, the latency of the N1 elicited by nontarget deviants was compared with that of the N1 elicited by the standards, measured at the electrode of maximal amplitude (O2). Peak latencies were submitted to a 2 within-subject \times 2 between-subject ANOVA with deviancy (standard/deviant) as a within-subject factor and language group (Spanish/English) as a between-subject factor.

Because some native speakers of Spanish were also Spanish–Catalan bilinguals, we investigated potential differences in attention allocation between groups by comparing ERPs elicited by mug standards and bowl targets on the one hand and cup standards and bowl targets on the other hand, because these comparisons always involved objects that have different names in both of the languages. P1s and DRNs elicited by “cup,” “mug,” and “bowl” (in identical time windows and the same electrodes as the analyses above) were subjected to repeated measures ANOVAs with object (cup–bowl/mug–bowl) as within-subject factor and language group (Spanish/English) as a between-subject factor. Because of the very high level of repetition involved in the oddball paradigm used here, we expected potential differences in attention to have a negligible impact on basic object discrimination as indexed by DRN. We therefore expected to find no interaction between object type and group in these comparisons.

RESULTS

Behavioral Data

Accuracy in the bowl detection task was above 90% in all participants and blocks, ($M_{\text{English}} = 0.94$, $SD = 0.02$; $M_{\text{Spanish}} = 0.93$, $SD = 0.02$). There was no significant differences between groups on target detection accuracy nor RTs ($t_1(25) = .62$, $p > .05$; $t_2(25) = .29$, $p > .05$).

Electrophysiological Data

Critical Comparison: Standard (Cup/Mug) versus Passive Deviant (Cup/Mug)

As expected, nontarget deviants elicited a greater DRN as compared with standards. This difference was qualified by a significant main effect of deviancy ($F(1, 25) = 10.3$, $p < .05$, $\eta_p^2 = 0.29$) with deviant stimuli eliciting more negative amplitudes than standard stimuli in the DRN window. The effect of deviancy further interacted with

language group ($F(1, 25) = 4.9$, $p < .05$, $\eta_p^2 = 0.16$), such that the deviancy effect was of significantly greater magnitude in English than Spanish participants (Figure 1A and B).

Post hoc test showed that there was no significant DRN effect in the Spanish group ($F(1, 12) = .46$, $p > .05$, $\eta_p^2 = 0.04$) but a significant effect in the English group ($F(1, 13) = 16.31$, $p = .001$, $\eta_p^2 = 0.56$; Figure 1C). Furthermore, there was no significant difference between standard and deviant conditions at any point in time in the DRN window in the Spanish participants, but standard and deviant conditions differed significantly from 135 to 177 msec in the English group (lower part of Figure 1A and B). To reduce the risk of type I errors and given the high levels of autocorrelation of ERP time series, we followed the method advocated by Guthrie and Buchwald (1991) where only sequences with a minimum of 12 consecutive significant t tests were considered (see, for instance, Kuipers & Thierry, 2011).

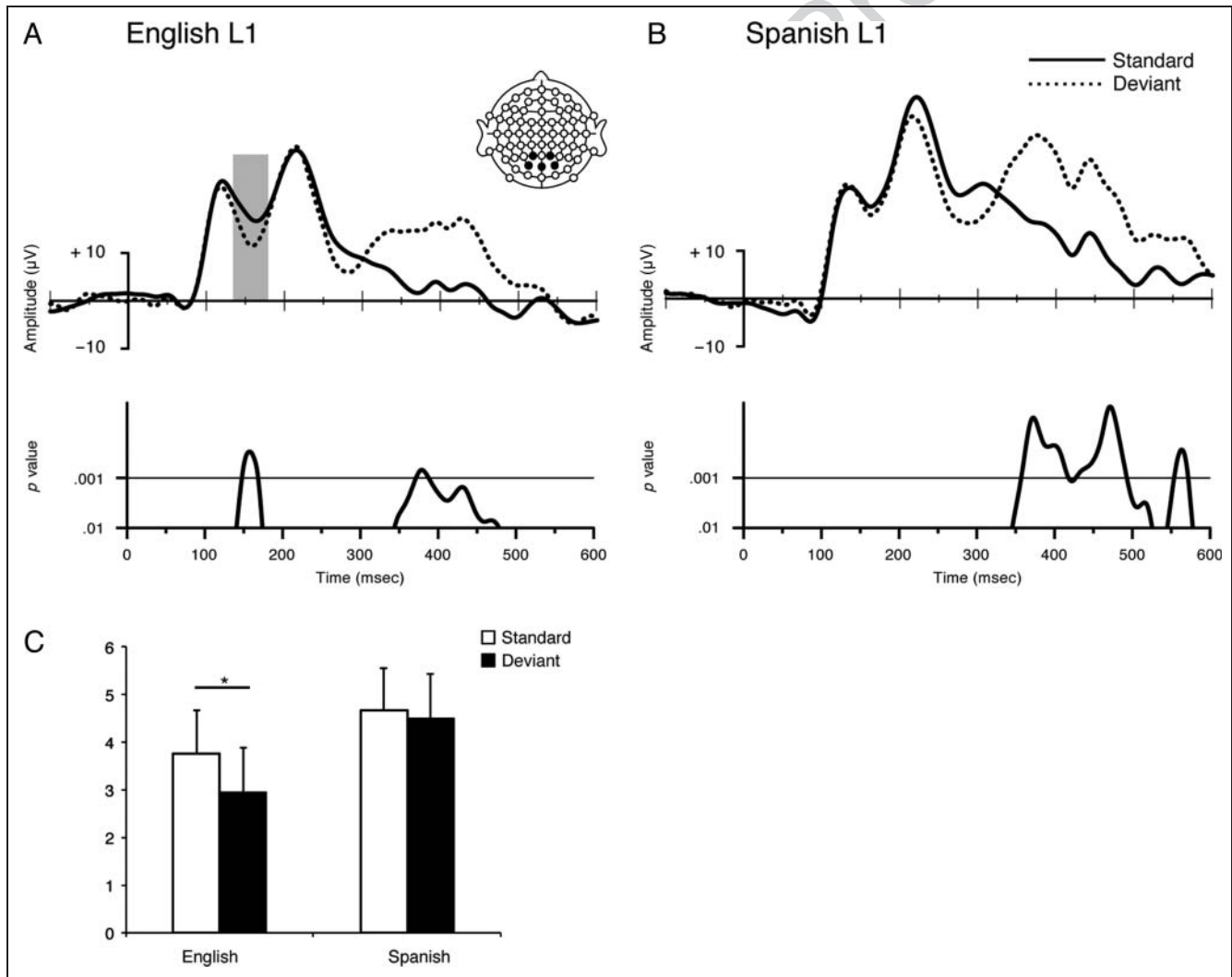
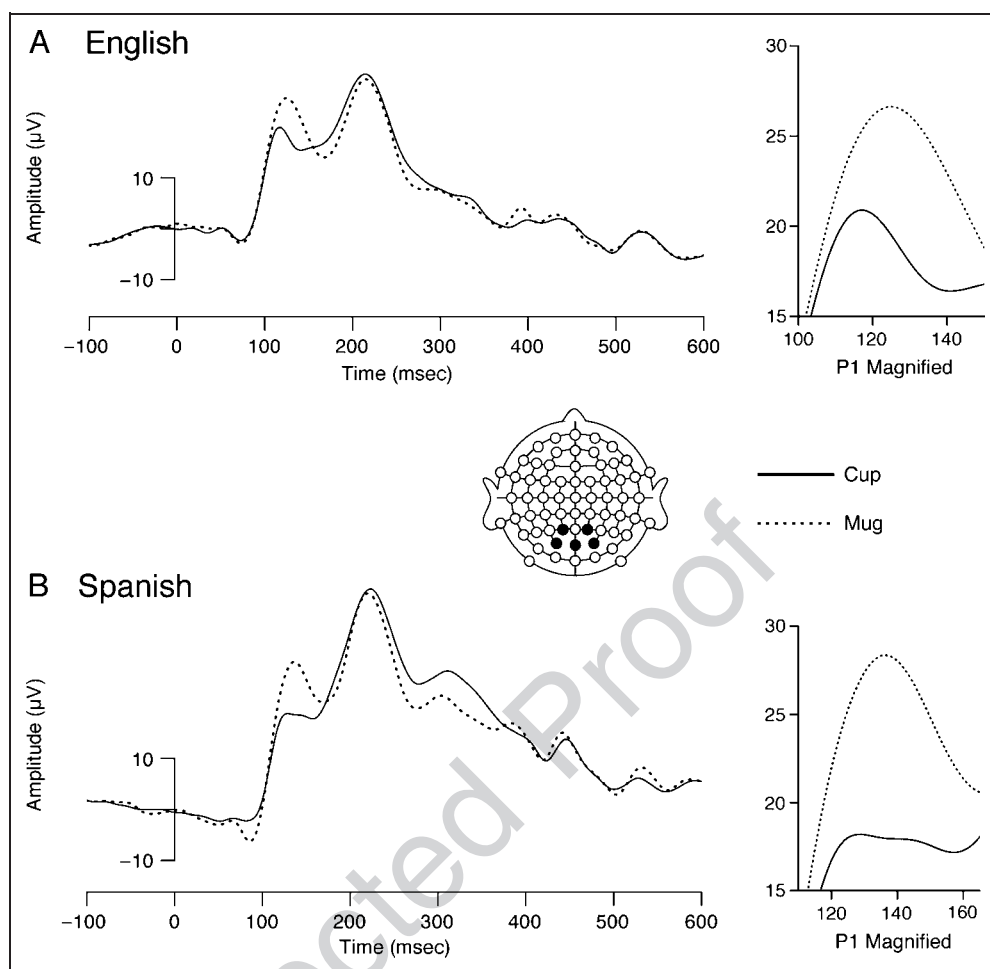


Figure 1. Event-related brain potentials elicited by standard and deviant stimuli averaged across blocks. ERPs and plots of p value of differences between conditions in (A) native speakers of English and (B) native speakers of Spanish. (C) Plot of DRN mean amplitude. Waveforms correspond to linear derivation of electrodes PO1, PO2, O1, OZ, and O2. Error bars depict SEM.

Figure 2. Event-related brain potentials elicited by cup and mug standards in each of the two experimental blocks. (A) Native speakers of English and (B) native speakers of Spanish. Waveforms correspond to linear derivation of electrodes PO1, PO2, O1, OZ, and O2.



Latency analyses of the DRN revealed no significant differences between group or condition in the window of interest ($F(1, 24) = 1.53, p > .05, \eta_p^2 = 0.06$).

ERPs elicited by standard stimuli in each of the two blocks considered separately (Figure 2) displayed significant differences in P1 mean amplitude (F_1) and latency (F_2) between cup and mug ($F_1(1, 24) = 5.76, p < .05, \eta_p^2 = 0.19$; $F_2(1, 24) = 17.56, p < .001, \eta_p^2 = 0.42$). Critically, these effects did not interact with participant group ($F_1(1, 24) = 1.29, p > .05, \eta_p^2 = 0.05$; $F_2(1, 24) = 3.2, p > .05, \eta_p^2 = 0.12$).

Control Comparison: Standard (Cup/Mug) versus Target (Bowl)

ANOVAs on the P1 revealed a significant effect of object type in both the mug versus bowl comparison ($F(1, 25) = 50.32, p < .0001, \eta_p^2 = 0.69$) and the cup versus bowl comparison ($F(1, 25) = 40.28, p < .0001, \eta_p^2 = 0.62$). Critically, there was no interaction between language group and object type in either comparisons (both $ps > .1$).

ANOVAs on the DRN revealed significant effect of object type in both the mug versus bowl comparison ($F(1, 25) = 40.28, p < .0001, \eta_p^2 = 0.62$) and the cup versus bowl

comparison ($F(1, 25) = 48.57, p < .0001, \eta_p^2 = 0.66$). Again, there was no interaction between language group and object type in either comparisons (both $ps > .1$).

DISCUSSION AND CONCLUSIONS

This study tested potential effects of language-specific terminology on early stages of visual perception and categorization based on the analysis of spontaneous modulations of the P1/N1 event-related brain potential complex. In a design controlling for perceptual features of the objects presented, ERPs successfully distinguished standards and deviants within the N1 range in native speakers of English but not in speakers of Spanish who name both these objects using the same noun. Moreover, when comparing the P1 elicited by the two objects presented as standards in each of the blocks, ERP differences were indistinguishable between groups.

The N1 range of ERPs is thought to index stages of visual processing beyond categorical discrimination (Dering, Martin, Moro, Pegna, & Thierry, 2011; Thierry, Martin, Downing, & Pegna, 2007a). Indeed, categorical effects have been reported in the domain of face processing in the P1 range and even earlier (Thierry et al., 2007a;

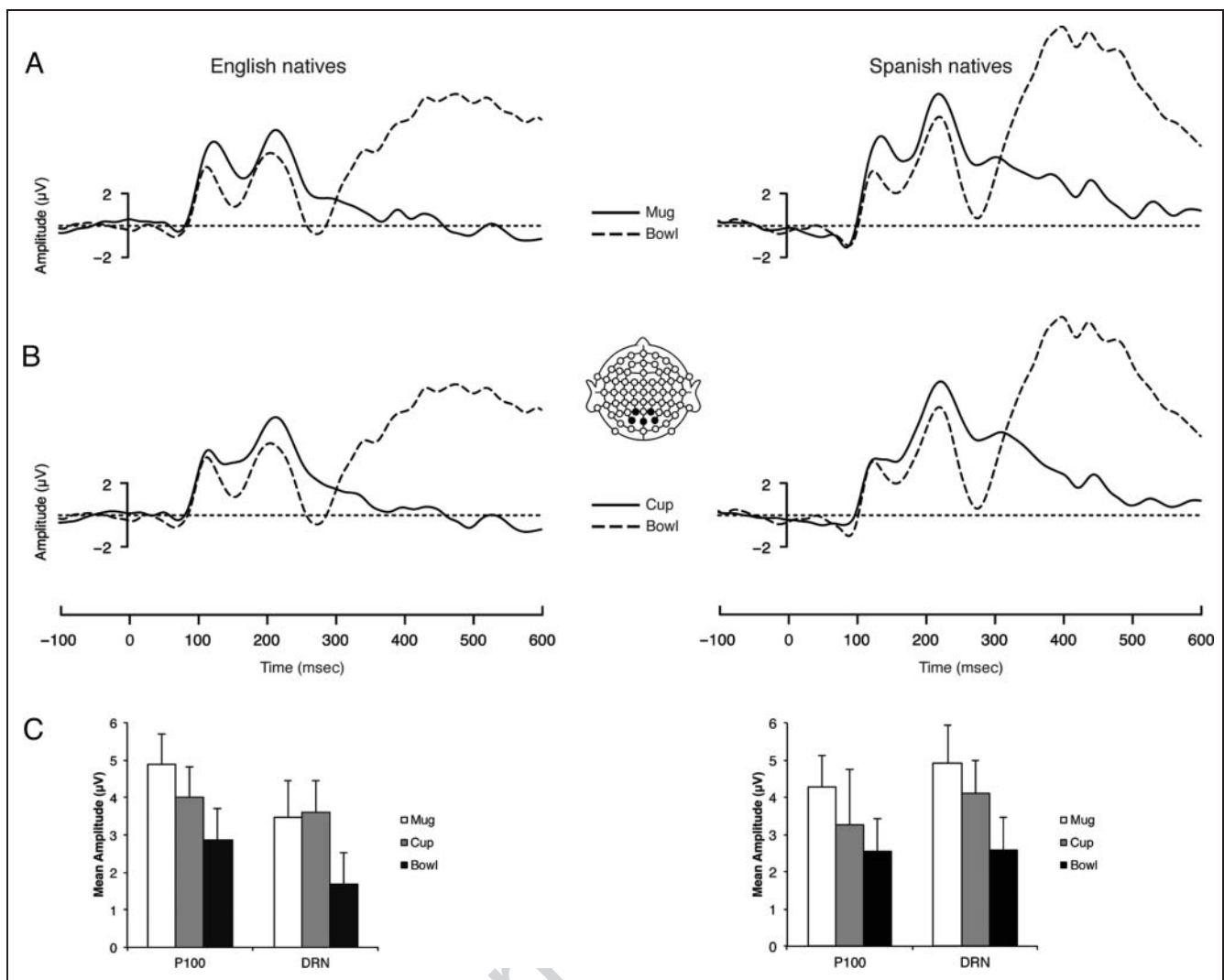


Figure 3. Event-related brain potentials elicited by (A) mug standards and bowl targets and (B) cup standards and bowl targets. (C) Plot of P1 and DRN mean amplitudes in both participant groups. Waveforms correspond to linear derivation of electrodes PO1, PO2, O1, OZ, and O2. Error bars depict *SEM*.

Thierry, Martin, Downing, & Pegna, 2007b; Seeck et al., 1997, 2001). Therefore, because it occurs beyond the P1 range, the DRN effect found here concerns relatively sophisticated levels of visual object processing—probably relating to object identity resolution. Critically, however, the DRN occurred before the temporal window in which lexical representation are considered to be accessed. Indeed, during practiced picture naming, Strijkers, Costa, and Thierry (2010) and Costa, Strijkers, Martin, and Thierry (2009) have established that lexical access occurs between 180 and 200 msec after picture onset. Here, significant differences were observed as early as 145 msec after picture onset. In addition, as shown by Strijkers and colleagues (Strijkers, Holcomb, & Costa, 2011), lexical access appears to be substantially delayed until ~350 msec after stimulus onset when there is no requirement to name the pictures (see also Blackford, Holcomb, Grainger, & Kuperberg, 2012). This was indeed the case here because participants were asked to press a button when they saw a specific

object and not instructed to name them. Thus, the influence of language-specific terminology on object processing does not merely result from online interaction with processes underlying lexical access. In other words, our finding is not simply an effect of language on language.

We report the N1 modulation recorded here as a DRN rather than a vMMN (the visual counterpart of the auditory MMN; Winkler et al., 2005; Czigler et al., 2002) because the vMMN proper is supposedly only elicited by visual stimuli presented outside the focus of attention, for example, in peripheral vision rather than fixation (Clifford et al., 2010). However, (a) the latency of the DRN effect we reported here is similar to that previously reported in vMMN studies (Pazo-Alvarez, Cadaveira, & Amenedo, 2003); (b) like our effect, the vMMN has a parieto-occipital topography with a right hemispheric predominance. Because the DRN in this study (peak time: ~160 msec at electrode O2) peaked substantially earlier and was observed at a different scalp location than N2 modulations

elicited by overt cognitive control (Folstein & Van Petten, 2007), we interpret this effect as an index of automatic, preattentive, and crucially, prelexical cognitive mechanism (Strijkers et al., 2010, 2011; Costa et al., 2009).

The P1 results further suggest that Spanish and English participants perceptually discriminated cup and mug pictures in a similar fashion. These two objects are indeed ostensibly different, and P1 amplitude has been shown to distinguish different object types previously (Dering et al., 2011; e.g., Thierry et al., 2007a). Therefore, the DRN effect observed in the N1 window cannot be explained by differences arising at more elementary stages of perceptual analysis preceding the N1 window. Furthermore, we consider the absence of between-group differences in the P1 range to be of fundamental importance because they could be underpinned by differences in cultural background or ethnic origin or even genetic factors and would therefore invalidate our results as merely stemming from different perceptual grooming in different environments.

Differences between groups in the P1 range could have been expected because our group has already reported such differences in a previous study of color perception (Thierry et al., 2009). However, it must be noted that the relationship between color terminology and P1 measurement was not trivial in that it did not yield a P1 amplitude by language group interaction. Expecting a reduction or cancellation of P1 differences between cups and mugs in the Spanish participants here would assume that perceptual differences between a cup and a mug are even more subtle than perceptual differences between two neighboring shades of blue, which have been shown to occur between 100 and 200 msec after stimulus onset (Fonteneau & Davidoff, 2007). We contend that cups and mugs are more discriminable at a perceptual level (at least by shape, size, and luminance) than two discs of the same size and color saturation, differing exclusively by their relative luminance. For example, people will argue indefinitely about color names at the green–blue or the navy–indigo border, but the same individuals will hardly argue as to what differentiates a mug and a cup shape. Therefore, it is reasonable to assume that P1 differences indexing early perceptual distinctions should effectively discriminate cups and mugs in both groups but that orientation responses measured by the DRN would be selectively affected by language terminology.

The fact that differences occur only in the N1 range and based on standard–deviant comparisons is essential to demonstrate an effect of language terminology on high-level perceptual processing. Additionally, these differences arising beyond the P1 range are consistent with an interactional account of linguistic relativity effects (Lupyan, 2012) because basic perception need not be changed for such effects to arise.

Our experimental design also allowed us to investigate potential attentional differences between the Spanish–Catalan speakers and English monolinguals. Indeed, one could argue that the interaction on the DRN could be a

result of better inhibition/monitoring mechanisms in the bilinguals. As suggested by our results, this was not the case because, when the items both had a different label in Spanish and English, the DRN elicited between target and standard had the same magnitude in the two groups. If Spanish participants had different attentional skills, and if such skills were generically reflected in DRN modulation, we would have expected the interaction observed in the critical comparison (mug/cup) to carry over to the case of comparisons with the target (bowl).

To our knowledge, this is the first neurophysiological demonstration of a relationship between native language and spontaneous object identity discrimination during visual perception, which goes beyond the observation of overt effects on object categorization (Pavlenko & Malt, 2010; Ameel et al., 2005, 2009). Furthermore, these findings generalize the linguistic relativity effects previously reported in the case of color perception (Liu et al., 2010; Thierry et al., 2009; Franklin et al., 2008) to the domain of object identity processing (Gilbert et al., 2008; arguably affecting higher-level cognitive representations). Overall, our results are incompatible with the view that language is functionally encapsulated in the human brain and fundamentally independent of, for example, visual cognition (Fodor, 1975, 2008; Chomsky, 2000; Pinker, 1995). On the contrary, they support an interactive conceptualization of the brain where language is highly integrated and can modulate ongoing cognitive processes such as object categorization and perception (Lupyan, 2012). Future studies will determine whether the effects reported here are confined to interactions within the left hemisphere (Mo, Xu, Kay, & Tan, 2011; Regier & Kay, 2009; Franklin et al., 2008; Roberson et al., 2008) and the extent to which they are adaptable over time (Athanasopoulos et al., 2010).

UNCITED REFERENCE

Pinker, 2007

Acknowledgments

B. B. is funded by RES-592-28-0001, and G. T. is funded by the Economic and Social Research Council (RES-000-23-0095) and the European Research Council (ERC-StG-209704). We thank Albert Costa, Clara Martin, Xavier Garcia, and Cristina Baus for their assistance with data collection in Spanish speakers.

Reprint requests should be sent to Bastien Boutonnet, School of Psychology, Bangor University, Bangor, Gwynedd LL57 2AS, United Kingdom, or via e-mail: b.boutonnet@bangor.ac.uk.

REFERENCES

- Ameel, E., Malt, B. C., Storms, G., & Van Assche, F. (2009). Semantic convergence in the bilingual lexicon. *Journal of Memory and Language*, *60*, 270–290.
- Ameel, E., Storms, G., Malt, B. C., & Sloman, S. A. (2005). How bilinguals solve the naming problem. *Journal of Memory and Language*, *53*, 60–80.

- Athanasopoulos, P. (2009). Cognitive representation of colour in bilinguals: The case of Greek blues. *Bilingualism: Language and Cognition*, *12*, 83.
- Athanasopoulos, P., Dering, B., Wiggert, A., Kuipers, J.-R., & Thierry, G. (2010). Perceptual shift in bilingualism: Brain potentials reveal plasticity in pre-attentive colour perception. *Cognition*, *116*, 437–443.
- Bassetti, B. (2007). Bilingualism and thought: Grammatical gender and concepts of objects in Italian–German bilingual children. *International Journal of Bilingualism*, *11*, 251–273.
- Blackford, T., Holcomb, P. J., Grainger, J., & Kuperberg, G. R. (2012). A funny thing happened on the way to articulation: N400 attenuation despite behavioral interference in picture naming. *Cognition*, *123*, 84–99.
- Boroditsky, L., Schmidt, L. A., & Phillips, W. (2003). Sex, syntax, and semantics. In D. Gentner & S. Goldin-Meadow, *Language in Mind: Advances in the Study of Language and Cognition* (pp. 61–79). Cambridge: MIT Press.
- Boutonnet, B., Athanasopoulos, P., & Thierry, G. (2012). Unconscious effects of grammatical gender during object categorisation. *Brain Research*, *1479*, 72–79.
- Bowerman, M., & Choi, S. (1991). Learning to express motion events in English and Korean: The influence of language-specific lexicalization patterns. *Cognition*, *41*, 83–121.
- Chomsky, N. (2000). *New horizons in the study of language and mind*. Cambridge: Cambridge University Press.
- Clifford, A., Holmes, A., Davies, I. R. L., & Franklin, A. (2010). Color categories affect pre-attentive color perception. *Biological Psychology*, *85*, 275–282.
- Costa, A., Strijkers, K., Martin, C., & Thierry, G. (2009). The time course of word retrieval revealed by event-related brain potentials during overt speech. *Proceedings of the National Academy of Sciences, U.S.A.*, *106*, 21442–21446.
- Csibra, G., Czigler, I., & Ambro, A. (1994). Effects of stimulus alternation, repetition and response requirements on event-related potentials to patterned visual stimuli. *Biological Psychology*, *37*, 115–132.
- Cubelli, R., Paolieri, D., Lotto, L., & Job, R. (2011). The effect of grammatical gender on object categorization. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *37*, 449–460.
- Czigler, I., Balázs, L., & Pató, L. G. (2004). Visual change detection: Event-related potentials are dependent on stimulus location in humans. *Neuroscience Letters*, *364*, 149–153.
- Czigler, I., Balázs, L., & Winkler, I. (2002). Memory-based detection of task-irrelevant visual changes. *Psychophysiology*, *39*, 869–873.
- Dering, B., Martin, C. D., Moro, S., Pegna, A. J., & Thierry, G. (2011). Face-sensitive processes one hundred milliseconds after picture onset. *Frontiers in Human Neuroscience*, *5*, 93.
- Feigenson, L., Dehaene, S., & Spelke, E. (2004). Core systems of number. *Trends in Cognitive Sciences*, *8*, 307–314.
- Flecken, M. (2010). Event conceptualization by early Dutch–German bilinguals: Insights from linguistic and eye-tracking data. *Bilingualism: Language and Cognition*, *14*, 61–77.
- Fodor, J. A. (1975). *The language of thought*. Cambridge: Harvard University Press.
- Fodor, J. A. (2008). *LOT 2: The language of thought revisited*. Oxford, UK: Clarendon Press.
- Folstein, J. R., & Van Petten, C. (2007). Influence of cognitive control and mismatch on the N2 component of the ERP: A review. *Psychophysiology*, *45*, 152–170.
- Fonteneau, E., & Davidoff, J. (2007). Neural correlates of colour categories. *NeuroReport*, *18*, 1323–1327.
- Forbes, J. N., Poulin-Dubois, D., Rivero, M. R., & Sera, M. D. (2008). Grammatical gender affects bilinguals' conceptual gender: Implications for linguistic relativity and decision making. *Open Applied Linguistics Journal*, *1*, 68–76.
- Franklin, A., Drivonikou, G., Clifford, A., Kay, P., Regier, T., & Davies, I. (2008). Lateralization of categorical perception of color changes with color term acquisition. *Proceedings of the National Academy of Sciences, U.S.A.*, *105*, 18221.
- Gallistel, C. R. (1989). Animal cognition: The representation of space, time and number. *Psychology*, *40*, 155–189.
- Gentner, D., & Goldin-Meadow, S. (2003). *Language in mind: Advances in the study of language and thought*. Cambridge, MA: MIT Press.
- Gilbert, A. L., Regier, T., Kay, P., & Ivry, R. B. (2006). Whorf hypothesis is supported in the right visual field but not the left. *Proceedings of the National Academy of Sciences, U.S.A.*, *103*, 489–494.
- Gilbert, A. L., Regier, T., Kay, P., & Ivry, R. B. (2008). Support for lateralization of the Whorf effect beyond the realm of color discrimination. *Brain and Language*, *105*, 91–98.
- Gordon, P. (2004). Numerical cognition without words: Evidence from Amazonia. *Science*, *306*, 496–499.
- Gratton, G., Coles, M. G. H., & Donchin, E. (1983). A new method for off-line removal of ocular artifact. *Electroencephalography and Clinical Neurophysiology*, *55*, 468–484.
- Gumperz, J. J., & Levinson, S. C. (1996). *Rethinking linguistic relativity*. Cambridge: Cambridge University Press.
- Guthrie, D., & Buchwald, J. S. (1991). Significance testing of difference potentials. *Psychophysiology*, *28*, 240–244.
- Klemfuss, N., Prinzmetal, W., & Ivry, R. B. (2012). How does language change perception: A cautionary note. *Frontiers in Psychology*, *3*, 78.
- Kuipers, J.-R., & Thierry, G. (2011). N400 amplitude reduction correlates with an increase in pupil size. *Frontiers in Human Neuroscience*, *5*, 61.
- Kurinski, E., & Sera, M. D. (2010). Does learning Spanish grammatical gender change English-speaking adults' categorization of inanimate objects? *Bilingualism: Language and Cognition*, *14*, 203–220.
- Lamme, V. A., & Roelfsema, P. R. (2000). The distinct modes of vision offered by feedforward and recurrent processing. *Trends in Neurosciences*, *23*, 571–579.
- Liu, Q., Li, H., Campos, J. L., Teeter, C., Tao, W., Zhang, Q., et al. (2010). Language suppression effects on the categorical perception of colour as evidenced through ERPs. *Biological Psychology*, *85*, 45–52.
- Lucy, J. A. (1992a). *Language diversity and thought: A reformulation of the linguistic relativity hypothesis*. Cambridge: Cambridge University Press.
- Lucy, J. A. (1992b). *Grammatical categories and cognition: A case study of the linguistic relativity hypothesis*. Cambridge: Cambridge University Press.
- Lupyan, G. (2012). Linguistically modulated perception and cognition: The label–feedback hypothesis. *Frontiers in Psychology*, *3*, 54.
- Majid, A., Bowerman, M., Kita, S., Haun, D. B. M., & Levinson, S. C. (2004). Can language restructure cognition? The case for space. *Trends in Cognitive Sciences*, *8*, 108–114.
- McClelland, J. L., & Rumelhart, D. E. (1981). An interactive activation model of context effects in letter perception: I. An account of basic findings. *Psychological Review*, *88*, 375.
- Mo, L., Xu, G., Kay, P., & Tan, L. H. (2011). Electrophysiological evidence for the left-lateralized effect of language on preattentive categorical perception of color. *Proceedings of the National Academy of Sciences, U.S.A.*, *108*, 14026–14030.
- Ozgen, E. (2004). Language, learning, and color perception. *Current Directions in Psychological Science*, *13*, 95–98.
- Papafragou, A., & Selimis, S. (2010). Event categorisation and language: A cross-linguistic study of motion. *Language and Cognitive Processes*, *25*, 224–260.

- Pavlenko, A., & Malt, B. C. (2010). Kitchen Russian: Cross-linguistic differences and first-language object naming by Russian–English bilinguals. *Bilingualism: Language and Cognition*, *14*, 19–45.
- Pazo-Alvarez, P., Cadaveira, F., & Amenedo, E. (2003). MMN in the visual modality: A review. *Biological Psychology*, *63*, 199–236.
- Phillips, W., & Boroditsky, L. (2003). Can quirks of grammar affect the way you think? Grammatical gender and object concepts. *Proceedings of the 25th Annual Meeting of the Cognitive Science Society*, 928–933.
- Picton, T., Bentin, S., Berg, P., Donchin, E., Hillyard, S., Johnson, R., Jr., et al. (2000). Guidelines for using human event-related potentials to study cognition: Recording standards and publication criteria. *Psychophysiology*, *37*, 127–152.
- Pinker, S. (1995). *The language instinct: The new science of language and mind*. London: Penguin.
- Pinker, S. (2007). *The stuff of thought: Language as a window into human nature*. New York: Viking.
- Regier, T., & Kay, P. (2009). Language, thought, and color: Whorf was half right. *Trends in Cognitive Sciences*, *13*, 439–446.
- Roberson, D., Davidoff, J., Davies, I. R. L., & Shapiro, L. R. (2005). Color categories: Evidence for the cultural relativity hypothesis. *Cognitive Psychology*, *50*, 378–411.
- Roberson, D., Pak, H., & Hanley, J. R. (2008). Categorical perception of colour in the left and right visual field is verbally mediated: Evidence from Korean. *Cognition*, *107*, 752–762.
- Saalbach, H., & Imai, M. (2007). Scope of linguistic influence: Does a classifier system alter object concepts? *Journal of Experimental Psychology: General*, *136*, 485–501.
- Seeck, M., Michel, C. M., Blanke, O., Thut, G., Landis, T., & Schomer, D. L. (2001). Intracranial neurophysiological correlates related to the processing of faces. *Epilepsy & Behavior*, *2*, 545–557.
- Seeck, M., Michel, C. M., Mainwaring, N., Cosgrove, R., Blume, H., Ives, J., et al. (1997). Evidence for rapid face recognition from human scalp and intracranial electrodes. *NeuroReport*, *8*, 2749–2754.
- Sera, M. D., Elieff, C., Forbes, J., Burch, M. C., Rodríguez, W., & Dubois, D. P. (2002). When language affects cognition and when it does not: An analysis of grammatical gender and classification. *Journal of Experimental Psychology: General*, *131*, 377–397.
- Strijkers, K., Costa, A., & Thierry, G. (2010). Tracking lexical access in speech production: Electrophysiological correlates of word frequency and cognate effects. *Cerebral Cortex*, *20*, 912–928.
- Strijkers, K., Holcomb, P. J., & Costa, A. (2011). Conscious intention to speak proactively facilitates lexical access during overt object naming. *Journal of Memory and Language*, *65*, 345–362.
- 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, U.S.A.*, *106*, 4567–4570.
- Thierry, G., Martin, C. D., Downing, P., & Pegna, A. J. (2007a). Controlling for interstimulus perceptual variance abolishes N170 face selectivity. *Nature Neuroscience*, *10*, 505–511.
- Thierry, G., Martin, C., Downing, P., & Pegna, A. (2007b). Is the N170 sensitive to the human face or to several intertwined perceptual and conceptual factors? *Nature Neuroscience*, *10*, 802–803.
- Turatto, M., Angrilli, A., Mazza, V., Umiltà, C., & Driver, J. (2002). Looking without seeing the background change: Electrophysiological correlates of change detection versus change blindness. *Cognition*, *84*, B1–B10.
- Whorf, B. L., Carroll, J. B., & Chase, S. (1956). *Language, thought and reality*. New York: Wiley.
- Winkler, I., Czigler, I., Sussman, E., Horváth, J., & Balázs, L. (2005). Preattentive binding of auditory and visual stimulus features. *Journal of Cognitive Neuroscience*, *17*, 320–339.
- Zhang, S., & Schmitt, B. (1998). Language-dependent classification: The mental representation of classifiers in cognition, memory, and ad evaluations. *Journal of Experimental Psychology: Applied*, *4*, 375.