Linguistic relativity and second language: How learning a second language may reshape cognition

Aina Casaponsa[1] and Guillaume Thierry[2,3]

[1] Department of Linguistics and English Language, Lancaster University, UK.

[2] School of Human and Behavioural Sciences, Bangor University, UK.

[3] Faculty of English, Adam Mickiewicz University, Poznan, Poland.

Orcid IDs:

- Aina Casaponsa: 0000-0002-8398-1435
- Guillaume Thierry: 0000-0003-1325-2656

Abstract

Language is the ultimate tool to describe our environment, communicate about our inner states, and accumulate knowledge that can be transferred between generations. Here, learning a new language is considered to offer new ways of describing, internalizing, and even perceiving the world. As learners accumulate experience with the second language (L2) for extended periods of time, their neurocognitive makeup evolves, like it does when learning new skills. Building on Benjamin Lee Whorf's and Edward Sapir's ideas introduced in the middle of the last century, the relationship between language and other cognitive processes is proposed to be dynamic in nature, following on logically from embodiment theory and the functional plasticity of the human brain. This chapter presents the theoretical basis for the idea that learning words and grammatical constructs can affect non-verbal cognition and reviews empirical evidence showcasing how learning an L2 can modulate the perception of colors, objects, and events.

Introduction and Definitions

Today, the world counts more than 7,000 languages, which can be grouped into approximately 150 families (Eberhard et al., 2019). Languages can vary considerably from one another across all linguistic representation levels (Evans & Levinson, 2009), including the number and nature of phonemes and graphemes, labels –auditory or written– used to classify objects or colors, and grammatical properties such as gender or aspect. The fact that each language has its own manifold characteristics and rules, which must be respected when we use it to communicate, has led linguists such as Sapir and Whorf (1956) to question whether individuals' behavior and their way of thinking could reflect the characteristics of their language.

The idea of an influence of language structure on thought has been extensively debated over the last century, and especially in the fields of linguistics, anthropology, philosophy, and more recently cognitive neuroscience (Athanasopoulos & Casaponsa, 2020; Everett, 2013; Lucy, 2016; Thierry, 2016). The Whorfian theory of linguistic relativity stems from the idea that the constant use of grammatical and structural characteristics idiosyncratic to a particular language orients attention and thereby "shapes the way" we perceive and think about the environment. As Whorf puts it: "users of markedly different grammars are pointed by their grammars toward different types of observations and different evaluations of externally similar acts of observation, and hence are not equivalent as observers but must arrive at somewhat different views of the world" (Whorf, 1956, p. 221). Sadly, Whorf died in his early forties, a few years after his mentor Sapir who also died prematurely at the age of 55. Neither of them had the chance to articulate their theoretical views in a way that would not leave the door wide open to interpretation. Whorf posthumously became the target of repeated, vitriolic attacks portraying him as a dilettante (Pullum, 1991), as irrational (Pinker, 2003), or even immoral (MacWhorter, 2014). However, character denigration is not a valid ground for dismissal and the debate should focus on formal scientific testing of the theory.

Our purpose here will not be to extensively discuss the validity of Whorf's views, which has received substantial empirical support in recent times (see Athanasopoulos & Casaponsa, 2020; Lupyan, 2012; Lupyan et al., 2020; Thierry, 2016), but rather to offer an account of how learning a second language (L2) may shift and reshape one's perception and conceptualization of the world. We therefore explore the idea that learning an additional language entails far more than the acquisition of language elements, structures and systems, but also new ways of thinking about and interacting with the world around us.

First, let us consider a few useful definitions:

- *Language*: Crystal and Robins (2018) define language as "a system of conventional spoken, manual (signed), or written symbols by means of which human beings, as members of a social group and participants in its culture, express themselves." The term 'conventional' implies some degree of arbitrariness: Language is usually not considered to make direct (analogic) references to the perceptual world, it uses arbitrary code making symbolic references to objects and concepts.
- *Native language (L1)*: The first and generally main language(s) of exposure from birth, that is, the language(s) one is first exposed to and starts learning from the environment during early development. Generally, a native language is the main language spoken by parents of a

child, but there can be many variations relating to language diversity in the environment, migration, childcare, or education.

- *Second language (L2)*: Language acquired by an individual after the foundations of the native language(s) have been laid (i.e., the case of sequential bilinguals). The term may also be used to describe the lesser used of two languages acquired simultaneously.
- *Bilingualism*: Broadly speaking, bilinguals are individuals who have a functional command of two languages.
- *Embodiment*: As Lakoff and Johnson (1999, pp. 41) put it: “Reason is not disembodied, as the tradition has largely held, but arises from the nature of our brains, bodies, and bodily experience”. In other words, the mind is built based on biophysiological, perceptual experiences of the world, rather than abstract representations detached from a physical input. This view, widely shared amongst the scientific community of cognitivists and neuroscientists, is often referred to as cognitive embodiment (see also Zappa & Frenck-Mestre, this volume).
- *Linguistic relativity*: The linguistic relativity hypothesis, also known as Sapir-Whorf hypothesis, postulates a link between language structure (mostly at lexical and syntactic levels) and mental representations (mostly at perceptual, attentional, and conceptual levels). In its broadest sense, linguistic relativity is concerned with a possible shaping of perception and (nonverbal) conceptualization by language. Linguistic relativity must be distinguished from the more extreme account of language determinism, which argues that language forms and structures are the foundation of thought and perception and thus, that nonverbal conceptual representations in fact do not exist. The Sapir-Whorf hypothesis has recently regained interest in the wake of findings from cognitive neuroscience regarding the neurophysiological bases of language and embodied cognition (Li et al., 2019; Pulvermüller, 2005).
- *Event-related potentials (ERPs):* This method, derived from electroencephalography, allows the researcher to isolate a prototypical response of the brain (potentials) to a particular set of stimuli (events) presented in a time-controlled fashion (see Dickson & Pelzl, this volume).

Historical Perspective and Theoretical Accounts: Language, Thought and the Brain

The world’s linguistic diversity alluded to in the introduction (Eberhard et al., 2019; Evans & Levinson, 2009) makes it immediately obvious that objects and actions, whether concrete or abstract, and their properties, are denoted by different words or labels in different languages. Such cross-language variation concerns not only the nature of words qualifying objects and actions, such as color or manner of motion (e.g., nouns and adjectives) but also grammatical properties (e.g., gender and aspect), which can convey functional relationships between agents and objects as well as contextual information about space and time. Languages also vary considerably in terms of supra-lexical syntax, for instance in their use of connectives or word order, leading to markedly diverse sentence organization.

This simple observation has led academics to question whether people’s mental representations and behavior could reflect linguistic characteristics specific to their language. Beyond the philosophical debate triggered by the idea of a relationship between language and thought, scientific evidence in the field of cognitive neuroscience indicates that the terms and grammatical constructs we use to communicate have a measurable impact on non-linguistic cognitive processes such as color perception (Athanasopoulos et al., 2010; Mo et al., 2011; Thierry et al., 2009; Xia et al., 2019), object categorization (Boutonnet et al., 2012; Boutonnet et al., 2013, Boutonnet & Lupyan, 2015; Casaponsa et al., 2022), the passage of time (Boroditsky et al., 2011; Bylund & Athanasopoulos, 2017; Li et al., 2018; Li et al., 2019), action conceptualization (Athanasopoulos & Bylund, 2013; Athanasopoulos et al., 2015; Bylund et al., 2013; Flecken et al., 2015; Francken et al., 2015; Kersten et al., 2010), and emotional processing (Barrett et al., 2007; Jonczyk et al., 2016; Lindquist et al., 2006; Wu & Thierry, 2012). Therefore, understanding how the brain organizes and processes linguistic information and how this information is integrated with other non-linguistic processes is especially important if we want to measure the impact of learning a L2 and improve learning efficiency.

Starting from a Hebbian perspective on learning (Hebb, 1949) and the idea that the nervous system is highly adaptative, cognitive processes such as perception, memory, reasoning, decision making, and language must necessarily be distributed across interactive assemblies of neurons rather than packed into discrete functional modules. Through experience, interconnections and associations between neurons are constantly modified and updated, restructuring the way in which our brain processes and stores information. For example, when we learn how to play an instrument, such as the piano, distant motor, visual, and auditory areas of the cortex become functionally related through the association of pressing a given key with a finger, a visually perceived note, and the sound produced by the piano. With only five hours of training involving key presses and hearing the associated sound, neuronal restructuring means that hearing a particular note is enough to specifically elicit activation in the motor cortex controlling the fingers (Lahav et al., 2007). Similar mechanisms obtain when we learn a language, resulting in fast and resilient associations between the sound of words and their meaning, as well as sensory attributes intrinsically connected to the association through neural subnets (embodiment). For example, seeing the word ‘coffee’ activates brain areas involved in visual word form recognition and comprehension as well as sensory areas such as olfactory or visual cortices (Pulvermüller & Fadiga, 2010). In the same vein, verb forms such as ‘running’ will not only stimulate brain areas involved in language, but also areas of the sensorimotor cortex, providing some of the strongest empirical evidence in support of embodied cognition (Pulvermüller et al., 2001; Pulvermüller et al., 2005; see Zappa & Frenck-Mestre, this volume).

We should, therefore, abandon theoretical conceptualizations of language as an encapsulated ‘module,’ somewhat functionally independent from other cognitive systems. With Pülvermüller (2005) and Barsalou (2008), we take the view that lexical-grammatical subsystems and “non-verbal” cognitive systems are reciprocally and intimately connected in the human brain (Athanasopoulos & Casaponsa 2020; Thierry, 2016). This implies that mental representations are distributed across neural networks that implement both linguistic information and their sensorimotor attributes. Conceptual representations, then, could arise from synchronization of neural firing patterns within and between multimodal and multisensory cell assemblies. Here, we consider in some detail the extent to which connections between our native language and other cognitive representations are penetrable and modified when we learn a L2.

One notable attempt at modelling interactions between language and non-verbal representations in long-term memory is Lupyan’s (2012) label feedback hypothesis. Lupyan posits that word representations (or labels) are crystallized independently from the concepts they refer to during learning, and that links between conceptual representations and labels are progressively strengthened over time. Eventually, a new word becomes well connected with the concept it refers to such that label representations are systematically activated when a concept is accessed and that, in turn, the corresponding concept becomes more salient for perception and attention. This proposal does not differ significantly from more general connectionist models of language processing, although it focusses rather selectively on the lexical (word/label) level of representation and thus does not readily account for relativistic effects relating to grammatical properties such as gender or aspect, let alone supra-lexical syntax (Sato & Athanasopoulos, 2018). In fact, a more general model of language-cognition interaction based on a Hebbian conception of the human brain (Pulvermüller, 1999) makes predictions similar to those made by Lupyan's label-feedback hypothesis, albeit based on synchronized activity in Hebbian cell assemblies, some representing word forms and others conceptual associations, be they strongly embodied or not.

Evidence for Linguistic Relativity Effects in Second Language Learning

Much more so than L1 acquisition, L2 learning requires the integration of new representations into a network of pre-existing ones. When we learn a new language, we must acquire new phonological and orthographic rules, new word forms, new syntactic and grammatical patterns, but also negotiate how the new information will coexist alongside native language representations. For example, learning the grammatical aspect of actions in progress conveyed by the progressive or gerund in Spanish should be easier for a native speaker of English (which also has a progressive form) than for a native speaker of German (which does not). But we argue here that learning a L2 is also about learning a new way of perceiving and thinking. A German learner of Spanish needs to both acquire a new grammatical device (e.g., the progressive) and adopt the corresponding conceptualisation of action unfolding over time (see Bylund & Athanasopoulos, 2015; Flecken et al., 2015; Pavlenko, 2014). In this section, we review examples of cognitive processes affected by language and how L2 learners might change the way they conceptualize and perceive their environment as a consequence.

Words, Perception, and Categorization

Color Categories

Thierry et al. (2009) conducted one of the first studies looking into early and automatic effects of terminology on color perception. They tested English monolinguals and Greek-English bilinguals (all native speakers of Greek living in the United Kingdom) in an oddball paradigm known to elicit modulations of the visual mismatch negativity (vMMN), an index of pre-attentive change detection caused by a stimulus presented rarely (deviant) with a stream of stimuli presented frequently (standards). Participants had to report squares and ignore round shapes. Stimulus color, which was not mentioned and task-irrelevant, varied such that light blue or green was a deviant and dark blue or green was a standard (and reciprocally in another block). The authors predicted that the existence of two basic terms in Greek for light blue (*ghalazio*) and dark blue (*ble*) would lead to differential vMMN modulation in Greek between 100-200ms post stimulus onset. Indeed, compared to Greek-English bilinguals, the vMMN was smaller in English monolinguals, who only have one basic term for blue, while green contrasts elicited similar vMMN effects in the two groups (Greek has also one term for green, *prasino*). When the data was subsequently split into two groups based on the length of stay in the United Kingdom, short stay (< two years) bilinguals showed a stronger light blue/dark blue differentiation effect than long stay bilinguals (Athanasopoulos et al., 2010).

Categorical perception thus seems affected by distinctions made by language terminology, and it can be reshaped by language experience. Subsequent studies have investigated whether visual perceptual restructuring can be observed with short periods of vocabulary learning (e.g., Clifford et al., 2012; Zhong et al. 2015; Zhou et al., 2010). With only five half-hour sessions of associating words with new color subcategories, Kwok et al. (2011) detected structural changes in cortical areas underpinning color perception, the same areas shown to activate jointly with language networks during color and shape association tasks (Siok et al., 2009; Tan et al., 2008). Thus, minimal experience with new labels to refer to colors leads to quick changes in categorical perception at a neurophysiological level. This evidence stands in contrast with the study by Athanasopoulos et al. (2010) where changes in color perception only emerged after 2 years of L2 immersion. Note however, that the latter study concerned category fusion (e.g., from *ghalazio* and *ble* to 'blue') as opposed to category refinement (from one category in L1 to two via learning new labels). Indeed, beyond the fact that new information is generally learnt faster than old information is forgotten, especially in the case of long-established knowledge, it is also arguably more ecologically valid to learn a new (relevant) category fast than forget an old one.

Object Categories

Spanish has two terms to distinguish glasses with and without a stem, *copa* and *vaso*, respectively, while in other languages, such as English, both objects are usually designated as 'glass'. Reciprocally, the word *taza* in Spanish encompasses two categories of objects in English: cups and mugs. Given that categorizing objects in the environment is one of our most fundamental cognitive abilities (Lakoff et al., 1987), the link between terminology and object categorization seems obvious. Categorization is an effective way of relating new objects and experiences to previous ones and therefore not only to memorize them but also associate them with meaning. This process is mostly implicit and automatic. If we come across an animal standing on two legs with a beak and feathers, we would likely categorize it as a kind of 'bird.' If it is bigger than a cat and its feathers are colorful, we might categorize it as 'pheasant' or 'peacock,' even if we are unsure what such birds look like. Similarly, when a native speaker of Spanish encounters a cup-looking object, they may assign it to the *taza* category, whereas a native speaker of English could assign the same object to the 'mug' category, based on properties such as size, proportions, thickness, material, et cetera. If systematic use of terms contributes to consolidating neural associations between mental representations of objects held in long-term memory, we can predict that objects assigned to these categories will be perceived as more different by a speaker of English than a speaker of Spanish.

Boutonnet et al. (2013) tested this prediction in an ERP oddball paradigm using the vMMN in a visual object recognition task. Rare presentations of a cup amongst frequent instances of a mug indeed failed to elicit, in native speakers of Spanish, the vMMN effect that was elicited by the same contrast in native speakers of English. To be clear, Spaniards could still tell the two objects apart –otherwise the results would support linguistic determinism– but their brain failed to meaningfully engage with the distinction during early perception. Consistent with the idea conveyed by Whorf, the authors argued that the two objects differently attract attention in speakers of English but not in native speakers of Spanish. What happens, then, when new categories are learnt, as is the case when the L2 makes categorical distinctions that do not exist in L1?

Maier et al. (2014) asked participants to learn new words to label invented objects (i.e., pseudo-objects). Some labels were associated with two different pseudo-objects, *de facto* belonging to the same category, while other labels were associated with only one object. Two days after training, which lasted only 45 minutes, P1 ERP modulations suggested that participants perceived objects belonging to the same category as more similar than objects belonging to a different category. Learning of new categories mediated by vocabulary thus modulates perceptual processing very quickly (see also Rabovsky et al., 2012; Yu et al., 2017a, 2017b). However, learning in this experiment was highly artificial and the objects had no pre-existing representation in memory or linguistic representation already associated with them, which is markedly different from the natural conditions in which an L2 is learned.

Pan and Jared (2020) tested the premises of the label-feedback hypothesis (Lupyan, 2012) in Chinese-English bilinguals using a vMMN paradigm with pictures of birds: a robin and an ostrich, whose Chinese labels both contain the character for bird, and a penguin and a pigeon, whose Chinese labels do not overlap. As predicted by the label-feedback hypothesis, the vMMN effect elicited by the robin/ostrich contrast was smaller than that elicited by the penguin/pigeon contrast, in a group of short-stay bilinguals (who had lived in Canada for <1.5 years). In a group of long-stay bilinguals (having lived in Canada for 2–9 years), the vMMN difference between the two contrasts tended to disappear, providing converging evidence with Athanasopoulos et al.'s (2010) observation in the domain of color.

Recently, our group conducted an adaptation of Boutonnet et al. (2013)'s ERP oddball paradigm in Spanish-English bilinguals using arrays of cups and mugs rather than single objects. Modulation of the vMMN by oddball arrays of peripherally presented objects, suggests that bilinguals (who learned L2 English at age 6 or later) become sensitive to category distinctions made in the L2 while retaining the distinctions made by the native language. We did not find any sign of category fusion, which is a tendency to perceive objects as more similar when a single term in L2 is associated with two categories dissociated in L1. Together the examples reviewed above point to fast and early effects of word learning on visual object processing, which suggests that L2 learning may have a general effect on perception in line with predictions from linguistic relativity (see Lucy, 2016). The next section reviews examples of linguistic relativity effects driven by the learning of L2 grammar.

Grammar and Conceptualization

Grammatical Gender and Object Conceptualization

Learning a L2 is far from limited to acquiring new vocabulary. Unless one has particularly poor verbatim memory, L2 grammatical properties and syntax are arguably more difficult to learn than new words, especially when such properties are absent in the native language (e.g., grammatical tense in Chinese) or when syntactic rules function differently in the two languages. Individuals whose native language does not have grammatical gender, for instance, find it particularly challenging when learning French, Spanish or German (Franceschina, 2005; Grüter et al., 2012; Sabourin & Stowe, 2008; see also Biondo et al., this volume, and Sabourin & Manning, this volume).

Contrary to conceptual gender (i.e., the gender of people and animals), grammatical gender of nouns is mostly arbitrary in nature and mappings between object and gender differ widely between languages. The French word *monde*, meaning 'world' is masculine (*le monde*), as it is in Spanish (*el mundo*), but it is feminine in German (*die Welt*). English learners of French or Spanish will need to not only learn the word for 'world' in the L2 but also its grammatical gender, which will dictate its form and agreement with other words in a sentence (determiner, verbs, adjectives, etc.). By contrast, German learners of French or Spanish will need to learn the word and the fact that it has a different gender in the L2.

One interesting question, then, given the obvious links between semantic and grammatical gender, is whether the constant use of male and female references when speaking of objects influences object conceptualization. Boutonnet et al. (2012) sought to determine whether Spanish-English bilinguals tested in a L2 English context would implicitly access the grammatical gender of objects presented to them as pictures (non-verbally), when they are engaged in a semantic categorization task. Participant were asked to indicate via button press whether the last of three objects presented in a row was from the same semantic category as the first two (as in, Tomato–Celery–Asparagus vs. Tomato–Celery–Truck). Unbeknownst to participants, the grammatical gender of the last object in Spanish was either consistent with that of the two first objects or not. In addition to expected modulations of the N400 by semantic priming (e.g., Kutas & Federmeier, 2011), bilinguals showed a Left Anterior Negativity (Friederici et al., 1993; see also Dickson & Pelzl, this volume) for objects with a grammatical gender inconsistent with that of the first two. Thus, Spanish-English bilinguals automatically, and presumably unconsciously, access grammatical gender information about everyday objects while immersed in an L2 context. The question then is whether they perceive concepts such as moon and sun in the same way as native speakers of English, given that Spanish grammar systematically associates moon with femininity (*la luna*) and the sun with masculinity (*el sol*).

The question of grammatical gender–semantic representation interactions has not yet been extensively addressed in L2 learners, possibly owing to the difficulty of controlling linguistic characteristics simultaneously in the two languages of a bilingual while measuring semantic interference of grammatical gender in a non-verbal task. Sato et al. (2020) sought to understand such interactions by measuring automatic responses of the brain to masculine and feminine faces primed by different sets of objects during a semantic categorization task. Objects were either stereotypically congruent or incongruent with the gender of the target face (e.g., a skirt or a tie). Unbeknown to participants, the authors also manipulated grammatical gender congruency in the participants' native language, French. Behaviorally, French-English bilinguals living in the UK and tested in English performed similarly as native English speakers, displaying also similar N300 modulations by conceptual gender priming (gender stereotypes). However, in contrast to English natives, bilingual participants implicitly matched the grammatical gender of the objects with the gender of faces, as shown by early perceptual modulations (indexed by N1 amplitude).

Thus, grammatical gender representations appear to be spontaneously accessed in a non-verbal context, which is likely to interfere with conceptual-semantic processing of visual objects.

Grammatical Aspect and Motion Conceptualization

Another grammatical characteristic that varies considerably across languages and which could affect perception and conceptualization is grammatical aspect. Athanasopoulos and Bylund (2013) tested the relationship between the characteristics of the native language and the perception of motion events in native speakers of English and native speakers of German. When German speakers describe an event depicted in a video clip (e.g., a woman walking towards a car), they tend to describe not only the action itself, but also the goal of the action (i.e., the fact that the woman walks *towards a car*), while English speakers tend to focus primarily on the action itself (i.e., the fact that the woman is *walking*). English and Spanish, contrary to German, have grammatical aspect. English, for instance, denotes unfolding actions using a gerund (e.g., children are *running*), whereas German lacks such progressive form. Athanasopoulos and Bylund (2013) confirmed their prediction that German speakers referred more to the goal of an action (e.g., children run *to school*) than English speakers when describing pictures. However, these observations were made in a context of language production (picture description) and could boil down to an effect of language-on-language, failing to capture how non-verbal representations are affected by linguistic constructs when language is not primarily involved (e.g., Thierry, 2016; Wessel-Tolvig & Paggio, 2016).

Nevertheless, Flecken et al. (2015) found converging evidence for effects of grammar on event conceptualization in a non-verbal, perceptual task. They used a P300 oddball paradigm, comparing brain responses to target pictures matching a preceding short animation in terms of either end point or trajectory in native English and native German speakers. The limitation mentioned above was also later addressed behaviorally using the ABX paradigm, in which participants are only required to compare events "visually" and decide which of two events (A/B) is most comparable to a third one (X), without resorting to a verbal description (e.g., Athanasopoulos et al., 2015). Athanasopoulos et al. (2015) not only showed greater focus on the goal of an unfolding action in native speakers of German compared to speakers of English, but also a shift from a German-like to an English-like pattern of behavior when bilingual participants switched from a verbal interference task conducted in English to one conducted in German. These results suggest that conceptualization depends on the language context and that shifts in conceptualization can be observed in real time, as the language context changes (see also Bylund & Athanasopoulos, 2015).

Other Factors to Consider

Together, the studies reviewed above establish that learning new lexical forms fosters neuroplasticity beyond linguistic representations. Such plasticity likely depends on several factors known to be involved in language learning: (a) developmental maturity and age of acquisition: the sooner we engage in learning, the greater the restructuring of the cognitive architecture (e.g., Kersten et al., 2010); (b) generic cognitive ability, such as intelligence, attention, and working memory (e.g., Casaponsa et al., 2015); and (c) the learning context, such as method, duration and rate of exposure (e.g., Athanasopoulos et al., 2015).

In addition to acquiring new language forms and structures, L2 learners have to manage constant and implicit activation of first language representations, which happen unconsciously and automatically (Thierry & Wu, 2007; Vaughan-Evans et al., 2015; Wu & Thierry, 2010), whether they are helpful, as in the case of cognates (Casaponsa et al., 2015) or counterproductive, as in the case of interlingual homographs or false friends. However, language co-activation is not always found, as has been observed in the case of affective words (Jończyk et al., 2016; Wu & Thierry, 2012), but also when equivalence across language involves complex mapping, as is the case of tense in Chinese learners of English (Li et al., 2018; Li et al., in press). In other words, the degree of correspondence between L1 and L2 in terms of, for example, translation equivalents and grammatical forms, may modulate the magnitude of linguistic relativity effects.

Finally, it is noteworthy that linguistic relativity likely concerns L2 learners to a greater degree than monolinguals, who probably seldom face the question of perceiving or conceptualizing the world under a 'different light'. For instance, a monolingual trainee might need to acquire extensive new vocabulary as they grow their expertise in a particular field of knowledge such as surgery or architecture, but any linguistic relativity effect will be restricted to the acquisition of new (open-class) words. It is highly unlikely that such individuals will learn new grammatical classes of (closed-class) words or novel syntactic patterns. Learning a second language, therefore, is likely to come hand-in-hand with stronger relativity effects, due to simultaneous learning of (many) new words together with new grammatical mechanisms and rules of syntax that shape the way we perceive and conceptualize our environment.

Innovations, Current Trends, and Future Directions

In the laboratory, participants can acquire new categories based on the acquisition of new words, via reorganization of links between cell assemblies dealing with perception and language representations. However, Athanasopoulos et al. (2010) showed in the case of color perception that category fusion does not seem to happen until two years of immersion in a foreign-speaking country have passed. Further investigation using objective measures of change detection at a perceptual and conceptual levels (e.g., mismatch negativity, P300, N400) is required that can measure longitudinally how category boundaries may be blurred over time as an individual learns a L2.

Preliminary evidence from our group suggests that expansion may not be measurable in the case of object categories, possibly because it is more informative to segment the world in as many categories as possible (category refinement), whereas regrouping previously learned categories under the same umbrella (category fusion) is arguably less useful. For instance, an English learner of Spanish will find it easier to learn the terms *copa* and *vaso* than the term *taza*. Indeed, learning to consider cups and mugs as belonging to one class of objects comes down to 'forgetting' the distinction between cups and mugs and, as argued throughout this chapter, this would require forgetting the distinction made in the native language between the words 'cup' and 'mug'. While this is theoretically possible, for instance in the case of severe attrition following a change of linguistic environment or language-selective aphasia, this should be observed in exceptional situations only and needs further investigation. This will require measures of implicit, unconscious information processing, using neuroscientific methods such as electrophysiology and magnetoencephalography, because participants and patients may not be able to overtly report perceptual differences, and such differences might not be detectable behaviorally either.

Such investigations could benefit from new methodological developments offering more sensitive and individually reliable measures, such as frequency tagging. Frequency tagging is an EEG technique capitalizing on brain entrainment to specific stimulus presentation frequencies. If two stimulus categories are visually and/or auditorily presented at two defined frequencies, different neural assemblies may latch on the two frequencies based on their "interest" in the particular stimuli, and thus start oscillating accordingly. While it is technically challenging to implement, this method has several potentially groundbreaking advantages: no task/response required, fast presentation rate meaning short experiment duration, statistical validity at the individual participant level, to name only a few. This latter advantage is crucial as it may allow correlating performance to individual proficiency in or exposure to L2.

Another area for future investigation concerns the speed at which the linguistic relativity effect can be "switched" on and off. To our knowledge, no immediate perceptual modulation by language context has yet been demonstrated in the literature. Future studies could test whether early, pre-attentive color perception can be modulated in real time by language of operation within the same L2 learners, possibly using oddball paradigm and the visual mismatch negativity as index combined with language cues, in order to set the language context in different blocks.

Further Reading

Comprehensive review of the Whorfian hypothesis on the intersection between language, cognition, and neuroscience:

Athanasopoulos, P., & Casaponsa, A. (2020). The Whorfian brain: Neuroscientific approaches to linguistic relativity. *Cognitive Neuropsychology, 37*(5-6), 393-412.

Theoretical review highlighting how linguistic labels can influence visual perception of the world around us:

Lupyan, G., Rahman, R. A., Boroditsky, L., & Clark, A. (2020). Effects of language on visual perception. *Trends in Cognitive Sciences*, *24*(11), 930-944.

First review emphasizing the need for neurolinguistic empirical evidence to support the Whorfian hypothesis:

Thierry, G. (2016). Neurolinguistic relativity: How language flexes human perception and cognition. *Language Learning, 66* (3): 690-713.

Acknowledgements

AC was supported by the British Academy/Leverhulme Small Research Grant (SG171343). G.T. was supported by the Polish National Agency for Academic Exchange (NAWA) under the NAWA Chair Program (PPN/PRO/2020/1/00006).

References

Athanasopoulos, P., & Bylund, E. (2013). Does grammatical aspect affect motion event cognition? A cross-linguistic comparison of English and Swedish speakers. *Cognitive Science*, *37*(2), 286-309. https://doi.org/10.1111/cogs.12006

Athanasopoulos, P., Bylund, E., Montero-Melis, G., Damjanovic, L., Schartner, A., Kibbe, A., Riches, N., & Thierry, G. (2015). Two languages, two minds: Flexible cognitive processing driven by language of operation. *Psychological Science*, *26*(4), 518-526. https://doi.org/10.1177/0956797614567509

Athanasopoulos, P., & Casaponsa, A. (2020). The Whorfian brain: Neuroscientific approaches to linguistic relativity. *Cognitive Neuropsychology, 37*(5-6), 393-412. https://doi.org/10.1080/02643294.2020.1769050

Athanasopoulos, P., Dering, B., Wiggett, A., Kuipers, J. R., & Thierry, G. (2010). Perceptual shift in bilingualism: Brain potentials reveal plasticity in pre-attentive colour perception. *Cognition, 116*(3), 437-443. https://doi.org/10.1016/j.cognition.2010.05.016

Barrett, L. F., Lindquist, K. A., & Gendron, M. (2007). Language as context for the perception of emotion. *Trends in Cognitive Sciences, 11*(8), 327-332. https://doi.org/10.1016/j.tics.2007.06.003

Barsalou, L. W. (2008). Cognitive and neural contributions to understanding the conceptual system. *Current Directions in Psychological Science*, *17*(2), 91-95. https://doi.org/10.1111/j.1467-8721.2008.00555.x

Biondo, N., Molinaro, N., & Mancini, S. (this volume). The neurolinguistics of the L2 morphological system: the role of grammar-related and speaker-related factors. In K. Morgan-Short and J. G. van Hell (Eds.), *The Routledge Handbook of Second Language Acquisition and Neurolinguistics*. Routledge.

Boroditsky, L., Fuhrman, O., & McCormick, K. (2011). Do English and Mandarin speakers think about time differently? *Cognition, 118*(1), 123-129. https://doi.org/10.1016/j.cognition.2010.09.010

Boutonnet, B., Athanasopoulos, P., & Thierry, G. (2012). Unconscious effects of grammatical gender during object categorisation. *Brain Research*, *1479*, 72-79. https://doi.org/10.1016/j.brainres.2012.08.044

Boutonnet, B., Dering, B., Viñas-Guasch, N., & Thierry, G. (2013). Seeing objects through the language glass. *Journal of Cognitive Neuroscience*, *25*(10), 1702-1710. https://doi.org/https://doi.org/10.1162/jocn_a_00415

Boutonnet, B., & Lupyan, G. (2015). Words jump-start vision: A label advantage in object recognition. *Journal of Neuroscience*, *35*(25), 9329-9335. https://doi.org/10.1523/JNEUROSCI.5111-14.2015

Bylund, E., & Athanasopoulos, P. (2015). Televised Whorf: Cognitive restructuring in advanced foreign language learners as a function of audiovisual media exposure. The *Modern Language Journal, 99*(S1), 123-137. https://doi.org/10.1111/j.1540-4781.2015.12182.x

Bylund, E., & Athanasopoulos, P. (2017). The Whorfian time warp: Representing duration through the language hourglass. *Journal of Experimental Psychology: General, 146*(7), 911. https://doi.org/https://doi.org/10.1037/xge0000314

Bylund, E., Athanasopoulos, P., & Oostendorp, M. (2013). Motion event cognition and grammatical aspect: Evidence from Afrikaans. *Linguistics*, *51*(5), 929-955. https://doi.org/10.1515/ling-2013-0033

Casaponsa, A., Antón, E., Pérez, A., & Duñabeitia, J. A. (2015). Foreign language comprehension achievement: Insights from the cognate facilitation effect. *Frontiers in psychology*, *6*, 588. https://doi.org/10.3389/fpsyg.2015.00588

Casaponsa, A., García-Guerrero, M. A., Martínez, A., Ojeda, N., Thierry, G., & Athanasopoulos, P. (2022). Electrophysiological evidence for a Whorfian double dissociation of categorical perception across two languages. PsyArXiv. https://doi.org/10.31234/osf.io/fdr3u

Clifford, A., Franklin, A., Holmes, A., Drivonikou, V. G., Özgen, E., & Davies, I. R. (2012). Neural correlates of acquired color category effects. *Brain and Cognition*, *80*(1), 126-143. https://doi.org/10.1016/j.bandc.2012.04.011

Crystal, D., & Robins, R. H. (2018). Language. In *Encyclopaedia britannica*. Retrieved from https://www.britannica.com/contributor/Robert-Henry-Robins/2482

Dickson, D., & Pelzl, E. (this volume). Using time-based electroencephalography to investigate L2. In K. Morgan-Short and J. G. van Hell (Eds.), *The Routledge Handbook of Second Language Acquisition and Neurolinguistics*. Routledge.

Eberhard, D. M., Simons, G. F., & Fennig, C. D. (2019). *Ethnologue: languages of the world*. Twenty-second edition. SIL International. http://www.ethnologue.com

Evans, N., & Levinson, S. C. (2009). The myth of language universals: language diversity and its importance for cognitive science. *Behavioral and Brain Science, 32*(5), 429-48. https://doi.org/10.1017/S0140525X0999094X

Everett, C. (2013). *Linguistic Relativity: Evidence across languages and cognitive domains*. De Gruyter Mouton. https://doi.irg/10.1515/9783110308143

Friederici, A. D., Pfeifer, E., & Hahne, A. (1993). Event-related brain potentials during natural speech processing: Effects of semantic, morphological and syntactic violations. *Cognitive Brain Research*, *1*(3), 183-192. https://doi.org/10.1016/0926-6410(93)90026-2

Flecken, M., Athanasopoulos, P., Kuipers, J. R., & Thierry, G. (2015). On the road to somewhere: Brain potentials reflect language effects on motion event perception. *Cognition*, *141*, 41-51. https://doi.org/10.1016/j.cognition.2015.04.006

Francken, J. C., Kok, P., Hagoort, P., & De Lange, F. P. (2015). The behavioral and neural effects of language on motion perception. *Journal of Cognitive Neuroscience*, *27*(1), 175-184. https://doi.org/10.1162/jocn_a_00682

Franceschina, F. (2005). *Fossilized second language grammars: The acquisition of grammatical gender*. John Benjamins Publishing. https://doi.org/10.1075/lald.38

Hebb, D. O. (1949). *The Organization of Behavior: A Psychological Theory*. Wiley & Sons.

Grüter, T., Lew-Williams, C., & Fernald, A. (2012). Grammatical gender in L2: A production or a real-time processing problem? *Second Language Research, 28*(2), 191-215. https://doi.org/10.1177/0267658312437

Jończyk, R., Boutonnet, B., Musiał, K., Hoemann, K., & Thierry, G. (2016). The bilingual brain turns a blind eye to negative statements in the second language. *Cognitive, Affective, & Behavioral Neuroscience*, *16*(3), 527-540. https://doi.org/10.3758/s13415-016-0411-x

Kersten, A. W., Meissner, C. A., Lechuga, J., Schwartz, B. L., Albrechtsen, J. S., & Iglesias, A. (2010). English speakers attend more strongly than Spanish speakers to manner of motion when classifying novel objects and events. *Journal of Experimental Psychology: General*, *139*(4), 638. https://doi.org/10.1037/a0020507

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*, 621. https://doi.org/10.1146/annurev.psych.093008.131123

Kwok, V., Niu, Z., Kay, P., Zhou, K., Mo, L., Jin, Z., ... & Tan, L. H. (2011). Learning new color names produces rapid increase in gray matter in the intact adult human cortex. *Proceedings of the National Academy of Sciences*, *108*(16), 6686-6688. https://doi.org/10.1073/pnas.1103217108

Lakoff, G. and Johnson, M. (1999). *Philosophy in the flesh: The embodied mind and its challenge to western thought.* Basic Books.

Lakoff, G. (1987). *Women, fire, & dangerous things: What categories reveal about the mind.* University of Chicago Press.

Lahav, A., Saltzman, E., & Schlaug, G. (2007). Action representation of sound: audiomotor recognition network while listening to newly acquired actions. *Journal of Neuroscience, 27*(2), 308-314. https://doi.org/10.1523/JNEUROSCI.4822-06.2007

Li, Y., Casaponsa, A., Wu, Y. J., & Thierry, G. (2019). Back to the future? How Chinese-English bilinguals switch between front and back orientation for time. *NeuroImage*, *203*, 116180. https://doi.org/10.1016/j.neuroimage.2019.116180

Li, Y., Casaponsa, A., Jones, M., & Thierry, G. (in press). Chinese learners of english are conceptually blind to temporal differences conveyed by tense. *Language Learning*

Li, Y., Jones, M., & Thierry, G. (2018). Timeline blurring in fluent Chinese-English bilinguals. *Brain Research*, *1701*, 93-102. https://doi.org/10.1016/j.brainres.2018.07.008

Lindquist, K. A., Barrett, L. F., Bliss-Moreau, E., & Russell, J.A. (2006). Language and the perception of emotion. *Emotion, 6*(1), 125-38. https://doi.org/10.1037/1528-3542.6.1.125

Lucy, J. (2016). The implications of linguistic relativity for language learning. In R. Alonso Alonso (Ed.), *Crosslinguistic Influence in Second Language Acquisition* (pp. 53-70). Multilingual Matters. https://doi.org/10.21832/9781783094837-006

Lupyan, G. (2012). Linguistically modulated perception and cognition: The label-feedback hypothesis. *Frontiers in Psychology*, *3*, 54. https://doi.org/10.3389/fpsyg.2012.00054

Lupyan, G., Rahman, R. A., Boroditsky, L., & Clark, A. (2020). Effects of language on visual perception. *Trends in Cognitive Sciences*, *24*(11), 930-944. https://doi.org/https://doi.org/10.1016/j.tics.2020.08.005

McWhorter, J. H. (2014). *The language hoax: Why the world looks the same in any language*. Oxford University Press.

Maier, M., Glage, P., Hohlfeld, A., & Abdel Rahman, R. (2014). Does the semantic content of verbal categories influence categorical perception? An ERP study. *Brain and Cognition, 91*, 1-10. https://doi.org/https://doi.org/10.1016/j.bandc.2014.07.008

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), 14026-30. https://doi.org/https://doi.org/10.1073/pnas.111186010

Pan, X., & Jared, D. (2021). Effects of Chinese word structure on object perception in Chinese–English bilinguals: Evidence from an ERP visual oddball paradigm. *Bilingualism: Language and Cognition, 24*(1), 111-123. https://doi.org/10.1017/S1366728920000206

Pavlenko, A. (2014). *The bilingual mind: And what it tells us about language and thought*. Cambridge University Press. https://doi.org/10.1017/CBO9781139021456

Pinker, S. (2003). *The language instinct: How the mind creates language*. Penguin.

Pullum, G. K. (1991). *The great Eskimo vocabulary hoax and other irreverent essays on the study of language*. University of Chicago Press.

Pulvermüller, F. (1999). Words in the brain's language. *Behavioral and brain sciences, 22*(2), 253-279.

Pulvermüller, F. (2005). Brain mechanisms linking language and action. *Nature Reviews Neuroscience*, *6*(7), 576-582. https://doi.org/10.1038/nrn1706

Pulvermüller, F., & Fadiga, L. (2010). Active perception: sensorimotor circuits as a cortical basis for language. *Nature Reviews Neuroscience 11*(5), 351. https://doi.org/10.1038/nrn2811

Pulvermüller, F., Härle, M., & Hummel, F. (2001). Walking or talking? Behavioral and neurophysiological correlates of action verb processing. *Brain and Language, 78*(2), 143-168. https://doi.org/10.1006/brln.2000.2390

Pulvermüller, F., Hauk, O., Nikulin, V. V., & Ilmoniemi, R. J. (2005). Functional links between motor and language systems. *European Journal of Neuroscience, 21*(3), 793-797. https://doi.org/10.1111/j.1460-9568.2005.03900.x

Rabovsky, M., Sommer, W., & Abdel Rahman, R. (2012). Depth of conceptual knowledge modulates visual processes during word reading. *Journal of Cognitive Neuroscience 24*(4), 990-1005. https://doi.org/10.1162/jocn_a_00117

Sabourin, L., & Manning, G. (this volume). Cross-linguistic transfer in N2 neurocognition. In K. Morgan-Short and J. G. van Hell (Eds.), *The Routledge Handbook of Second Language Acquisition and Neurolinguistics*. Routledge.

Sabourin, L., & Stowe., L. A. (2008). Second language processing: When are first and second languages processed similarly? *Second Language Research, 24*(3), 397-430. https://doi.org/10.1177/02676583080901

Sato, S., & Athanasopoulos, P. (2018). Grammatical gender affects gender perception: Evidence for the structural-feedback hypothesis. *Cognition, 176*, 220-231. https://doi.org/10.1016/j.cognition.2018.03.014

Sato, S., Casaponsa, A., & Athanasopoulos, P. (2020). Flexing gender perception: Brain potentials reveal the cognitive permeability of grammatical information. *Cognitive Science*, *44*(9), e12884. https://doi.org/10.1111/cogs.12884

Siok, W. T., Kay, P., Wang, W. S., Chan, A. H., Chen, L., Luke, K. K., & Tan, L. H. (2009). Language regions of brain are operative in color perception. *Proceedings of the National Academy of Sciences*, *106*(20), 8140. https://doi.org/10.1073/pnas.0903627106

Tan, L. H., Chan, A. H., Kay, P., Khong, P. L., Yip, L. K., & Luke, K. K. (2008). Language affects patterns of brain activation associated with perceptual decision. *Proceedings of the National Academy of Sciences*, *105*(10), 4004-4009. https://doi.org/10.1073/pnas.0800055105

Thierry, G. (2016). Neurolinguistic Relativity: How Language Flexes Human Perception and Cognition. *Language Learning, 66* (3), 690-713. https://doi.org/10.1111/lang.12186

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)*, 4567-70. https://doi.org/10.1073/pnas.0811155106

Vaughan-Evans, A., Thierry, G., Kuipers, J.-R., & Jones, M. W. (2014). Anomalous transfer of syntax between languages. *The Journal of Neuroscience*, *34*(24), 8333–8335. https://doi.org/10.1523/JNEUROSCI.0665-14.2014

Wessel-Tolvig, B., & Paggio, P. (2016). Revisiting the thinking-for-speaking hypothesis: Speech and gesture representation of motion in Danish and Italian. *Journal of Pragmatics*, *99*, 39-61. https://doi.org/10.1016/j.pragma.2016.05.004

Whorf, B. L. (1956). *Language, Thought, and Reality: Selected Writings of Benjamin Lee Whorf* (J. B. Carroll, Ed.). The MIT Press.

Wu, Y. J., & Thierry, G. (2010). Chinese–english bilinguals reading english hear chinese. *Journal of Neuroscience*, *30*(22), 7646-7651. https://doi.org/10.1523/JNEUROSCI.1602-10.2010

Wu, Y. J., & Thierry, G. (2012). How reading in a second language protects your heart. *Journal of Neuroscience*, *32*(19), 6485-6489. https://doi.org/10.1523/JNEUROSCI.6119-11.2012

Xia, T., Xu, G., & Mo, L. (2019). Bi-lateralized Whorfian effect in color perception: Evidence from Chinese Sign Language. *Journal of Neurolinguistics, 49,* 189-201. https://doi.org/10.1016/j.jneuroling.2018.07.004

Yu, M., Li, Y., Mo, C., & Mo, L. (2017a). Newly learned categories induce pre-attentive categorical perception of faces. *Scientific Reports, 7*(1), 14006. https://doi.org/10.1038/s41598-017-14104-6

Yu, M., Mo, C., Zeng, T., Zhao, S., & Mo, L. (2017b). Short-term trained lexical categories affect preattentive shape perception: Evidence from vMMN. *Psychophysiology, 54*(3), 462-468. https://doi.org/10.1111/psyp.12797

Zappa, A., & Frenck-Mestre, C. (this volume). Embodied L2 processing and learning from a neurocognitive perspective. In K. Morgan-Short and J. G. van Hell (Eds.), *The Routledge Handbook of Second Language Acquisition and Neurolinguistics*. Routledge.

Zhong, W., Li, Y., Li, P., Xu, G., & Mo, L. (2015). Short-term trained lexical categories produce preattentive categorical perception of color: Evidence from ERPs. *Psychophysiology*, *52*(1), 98–106. https://doi.org/10.1111/psyp.12294

Zhou, K., Mo, L., Kay, P., Kwok, V. P., Ip, T. N., & Tan, L. H. (2010). Newly trained lexical categories produce lateralized categorical perception of color. *Proceedings of the National Academy of Sciences*, *107*(22), 9974-9978. https://doi.org/10.1073/pnas.1005669107