MASTER IN COGNITIVE SCIENCE AND LANGUAGE

MASTER THESIS

November 2020

Towards a Theoretical Framework for Linguistic Relativity Research

by ANDONI FERNÁNDEZ PERALES

Under the supervision of:

MARC GUASCH MOIX

UNIVERSITAT ROVIRA i VIRGILI

Abstract

Linguistic relativity research is predicated on a basic premise: differences in language cause differences in cognition. However, the widely adopted paradigm set forth by Benjamin Lee Whorf has failed to provide satisfying evidence for it. In this paper, we identify the fundamental flaws of linguistic relativity as a scientific discipline as well as their historical causes. By doing so, we justify the elaboration of a theoretical framework that aims to instigate the creation of a cohesive and empirically valid paradigm by focusing on three objectives: first, clearly defining what constitutes a relativistic relationship between language and cognition; second, presenting a workable methodological process that aligns with that conception; and lastly, offering a formal classification of linguistic traits that allows for their organization as minimal, empirically-testable units, so they can better fit experimental activity.

Contents

1 Introduction

The notion of "linguistic relativity" rests on a fundamental premise: language influences the manner in which we experience the world. It is implied by this proposition, then, that how each individual perceives, processes and represents reality is *relative* to the different *linguistic* structures they interact with. This basic tenet, although it inaugurated a whole new line of academic research in the first half of the twentieth century, has been widely and unceasingly contested[1] since not long after it was first formulated by Benjamin Lee Whorf as the 'linguistic relativity principle' (1940a, 1940b), making it a controversial domain. John A. Lucy, one of the most prolific contemporary scholars in this field, asserted in this regard that "few ideas generate as much interest and controversy as the linguistic relativity hypothesis" (1997, p. 291).

But *why* is it that linguistic relativity gets so much attention? Why is it able to congregate both enthusiasts and opponents alike through decades-long debates? The answer, we reckon, might point towards the condition of human language as one of the deeply essential and distinctive components that define us as a species. Moreover, to the fact that our dependence on language is universal as well as constant: its familiarity, it seems, can often times serve as an invitation to reflect upon its nature, or in the very least as a license to discuss or form an opinion on it when prompted. Talking about language is, therefore, an accessible way of delving into such an interesting topic as the constitutive qualities of the human being—*what* we are. In fact, this very reason might explain why the role of language as the potential cause or reflection of our unique cognition—*why* we are what we are—is a primordial matter of intellectual discussion; so much so, that it "can probably be traced back to the dawn of philosophy" (Gumperz and Levinson, 1996, p. 3).

Admittedly, our reasoning can only account for the attention this intriguing line of inquiry has accrued over time; what about rejection and controversy, then? Negative reactions to linguistic relativity have usually centered around refuting its core proposition on the basis of its two most salient flaws: its ill-defined object of study, and the resulting lack of pertinent approaches to the rigorous examination of its substance. We shall see that these are directly responsible for the heavily critizised, "supposedly mind-broadening anecdotes [that] owe their appeal to a patronizing willingness to treat other cultures' psychologies as weird and exotic compared to our own" (Pinker, 1994, p. 64). However, we will argue in this paper that the thing that, for many, makes its *principle* a highly contestable *hypothesis* is not the elemental connection between language and cognition that it naturally stemmed from, but rather the particular paradigm that has been followed historically to illustrate it: a mostly unchanged set of defective theoretical presuppositions and methodologies inherited by the vast majority of literature in this field.

[1] Chomsky, Malotki (1983), and Pinker (1994) are some of its famous detractors.

These practices have almost exclusively focused on emphasizing semantic (*linguistic*) differences in the categorization of—chiefly—space, time, color, objects, quantities, and gender[2] across languages that, they claim, is significantly associated with variation in the general (*non-linguistic*) cognition of their speakers. For instance, if a given language conceptualizes time as a vertical line that progresses upwards, that could imply that its speakers perceive time differently than speakers of a language that conceptualizes it as a horizontal line. The problem with this approach is that it tries to prove general cognitive variation (differences in perception and processing of time, space, objects, color, etcetera) by simply describing different categorization strategies adopted by each language, instead of looking for structural and functional neural dissimilarities between individuals. Establishing that each language provides a unique conceptual model to their speakers is surely very far from proving that it has a deep and lasting effect on the foundation of their minds.

The first objective of the present study is to uncouple that particular perspective—what has come to be called the "Sapir-Whorf hypothesis" or "Whorfianism": the American paradigm set forth by Sapir, Whorf, Lenneberg, Brown, and so on—from what we propose should be understood by *linguistic relativity* as a subject and area of research in general terms—the systematic assessment of influencing linguistic factors on general cognition. In order to set clear boundaries between the underlying, primary idea of linguistic relativity, and this concrete academic tradition that revolves around it, we will go over various examples of other specific formulations, made by other authors long before Whorf, of the very same notion. In doing so, we'll claim that it's *not*, in fact, linguistic relativity's basic premise what has been contested so far, but the methods and rationale that have been commonly used to support it—all emerging from certain landmarks in its historical development that, we will maintain, are the cause of the characteristic indefiniteness and excessive reliance on semantics we just alluded to.

This preliminary contextual analysis appears to be absolutely necessary, as it will serve a very specific purpose: to validate the integrity of its core idea regardless of its practical treatment in the last eigthy years, providing us with a conceptual justification for the true goal and motivation behind this paper; namely, laying out a theoretical framework for linguistic relativity research. Drawing on revisions already suggested by other authors, our aim is to propose the implementation of a model for linguistic relativity that clearly defines its object of study, offering a formal structure intended to select and arrange *every* instance of linguistic phenomena beyond semantics that could have an influence on non-linguistic, cognitive traits, in the hopes that it can help start painting a complete and comprehensible picture of language's impact on the mind that can be effectively used as a guideline for future research. To our surprise, Lucy (2016) already attests to the

[2]Everett (2013, p. 62) presents these "cognitive or experiential domain(s)" as the more salient with regards to linguistic relativity research.

necessity of such an enterprise:

> Without some guiding sense of which aspects of language variation should matter most for thought and why, the field risks dissipating its empirical energies and failing to develop a coherent theoretical account of the relation of language variation to thought (p. 500).

We believe that only the design of a systematized, exhaustive framework of this sort can facilitate the development of a new paradigm in linguistic relativity research—one that, through an inevitable reliance on multidisciplinarity (philosophy, psychology, linguistics, neuroscience, etcetera), can allow for relevant progress in a domain that deals with such a profoundly important subject matter. What follows is an initial attempt at starting to put together such a framework.

2 The problem with linguistic relativity

The problematic nature of linguistic relativity has always been its own biggest hurdle. Acknowledging this fact begs the following question: what, *exactly*, is its main fault? Neither technical complexity nor a lack of resources can, within the bounds of reason, be deemed responsible for hindering its scientific advancement. Rather, its greatest impediments reside elsewhere: in a critical flaw of concept underpinning its inception, sustained through its entire development as a discipline. Precisely, it's how vaguely and inappropiately its object of study was defined what tends to disrupt its research, usually constraining it to the limited scope of semantic categorization variability between languages—the assessment of different mental representations induced by linguistically-articulated thought.

Of course, this approach is at odds with the theoretical starting point of the field, since it doesn't concern itself equally with the impact those linguistic phenomena exert on general cognition, concentrating instead on linguistic variability in certain semantic domains (time, motion, objects, color, and so on). Cognition, then, is reduced to the differing ways of symbolically encoding reality each language utilizes. On top of being redundant, as this kind of perspective doesn't really go beyond the examination of *purely* linguistic mechanisms (word-meaning pairings), it dilutes the broad relation between language and cognition that linguistic relativity should try to illustrate in a clear and exhaustive manner, instead of fostering research practices that are either theoretically incoherent with that purpose or so empirically misguided that they cannot possibly leave the realm of speculation.

But before we go into detail about the constitutive flaw of the field's academic tradition[3], we should pinpoint what we think are the exact *causes* of its occurrence, so we

[3]One could argue, as does Lucy (2016, p. 500), that it's actually "a scattered array of little traditions";

can better describe it later. Because we believe that they can be accurately located in time (Whorf's controversial remarks on Hopi language, in 1936, and his first explicit formulations of the linguistic relativity principle, in 1940; Brown and Lenneberg carrying out the first experimental work ever in this domain, in 1954, taking Whorf's writings as their point of reference), we'll also provide some historical context, briefly highlighting a few of the numerous instances where linguistic relativity's central idea was previously discussed or hinted at. This way, we intend to preserve the legitimacy of its principle by subtracting the faults of its later application, which we will eventually reach and give an account of.

2.1 Whorfianism is not linguistic relativity: a historical perspective

Disentangling *Whorfianism* from *linguistic relativity*[4] (both as conceived in this paper) doesn't strike us as a difficult task: although a complete genealogy of ideas linking language to cognition in a relativistic sense is yet to be published, there has been a tendency in this field to dig up old writings, as it were, in search of precursors to the paradigm set by Whorf and others, so we have many examples to draw from as a result.

Although Elffers (1996, p. 75) characterizes this fact in a negative light, as a "tendency of seeing relativism *wherever* some thought about language and thought is presented",[5] and while it is absolutely true that there have been several misconceptions in this respect, one can be equally sure that the basis of linguistic relativity has otherwise "deep historical roots" (Everett, 2013, p. 1), for even casual findings of authors touching upon the subject—in a very evident manner, that is—well before its birth as an area of study aren't rare. For instance, Spanish writer and philosopher Miguel de Unamuno, during the early years of the twentieth century, already dabbled with the idea that language could potentially have a transformative effect on the human mind, giving a possible justification to its emergence in evolutionary terms:

> Reason, that which we call reason—our reflective self-knowledge, the distinguishing quality of man—, is a social byproduct. It owes its origin, perhaps, to language. We think articulately—that is, reflectively—thanks to articulated language, and this language stemmed from the necessity of transmitting our thoughts to our fellow men (*Del sentimiento trágico de la vida*, 1913, p. 70, translation mine).

nonetheless, we believe that most of them are indelibly marked by the same fundamental problem stemming from linguistic relativity's original formulation, and so they pertain to the same school. We could posit, in any case, that they are "traditions within a tradition".

[4]Although it's closely tied to Whorfianism and they are usually conflated, we use and maintain the same term Benjamin Lee Whorf introduced, for convenience as well as for how accurately it represents the idea it denotes.

[5]Emphasis mine.

Another interesting case that, to our knowledge, hasn't been associated to linguistic relativity either, is that of Ancient Greek philosopher Democritus of Abdera (460 - 370 BC), one of the fathers of atomism. Perhaps "the first person known to have reflected upon language independently of all other existing objects" (Di Cesare apud Joseph, 2000, p. 14), his ideas look to be closely tied to our central subject:

> ...according to Democritus, linguistic utterances are *doubly* conventional: conventional means for denoting the elements of a sensory world which itself exists in the realm of convention.... the double conventionality does lead people to confuse *nomos* [custom, convention] for *physis* [nature, reality] when they mistake their linguistic conventions for reality. This concern is a binding thread among the Presocratic philosophers, as well as Socrates and Plato. The obfuscation of truth by language was at the center of their critique of language itself... The idea that language might operate strictly on the basis of convention seems not to have bothered Democritus, who stated it in its strongest form (ibidem).

Democritus' view of language here seems to be opposed to Aristotle's universalistic standpoint, as he claimed in *De Interpretatione* (16a, 3-8) that all possible symbolic referents are the *same* regardless of how they are expressed through each language:

> ...just as written marks are not the same for all men, neither are spoken sounds. But what these are in the first place signs of—affections of the soul—are the same for all; and what these affections are likenesess of—actual things—are also the same (apud De Cuypere & Willems, 2008, p. 308).

More interestingly, Democritus' position appears to positively match, even with great precision, Elffers' very own criteria to discriminate between what can "signify linguistic relativism" and what can't:

> It [linguistic relativity] presupposes *a non-universalistic view of concepts and grammar*. Therefore, universalistic views of concepts (like Platonism or Cartesianism) and universalistic views of grammar (like 17th-century General Grammar and its successors) preclude linguistic relativism (op. cit., p. 76, emphasis mine).

The philosopher's *double conventionality* seems to adhere to this description: (1) the conventional nature of the "sensory world", or perception, leads to a "non-universalistic view of concepts", and (2) the *strictly* conventional character of language reflects a "non-universalistic view of grammar".

Unlike Aristotle, Democritus appeared to believe that the objects of reality each language represents *aren't* the same; moreover, that the perception mediating between

reality and individuals—the speakers of those languages—*also* differs based upon convention. This is coherent with linguistic relativity arranging reality in diverse pre-linguistic concepts that are transformed into linguistic ones through the acquisition of a language, each one doing it in an essentially different way. A "universalistic view of grammar", on the other hand, would make every language denote the same basic linguistic concepts in different superficial forms; furthermore, a "universalistic view of concepts" would make these linguistically-articulated concepts capable of referring to the same ultimate reality. It's easy to see which side the Greek philosopher's theory would fall into.

According to our brief rundown, Democritus' apparently relativistic conception of language could very well be one of the *earliest* among them—it's a shame, nonetheless, that his books on language, as recorded by historian Diogenes Laertius (*On Homer or On the Property of Expression and Dialectal Locutions*, *On Words*, *On Poetry*, etcetera) were lost (Joseph, ibidem), seeing that they could have helped us explore his ideas in a more comprehensive fashion.

Other, more recent and considerably well-known examples include German philosopher Friedrich Nietzsche (1844-1900), whose observations on language we reproduce here might have caught Unamuno's attention, as they resemble his own analysis to a large extent:

> . . . consciousness has developed only under the pressure of the need for communication. . . . As the most endangered animal, he [the human being] *needed* help and protection, he needed his peers, he had to learn to express his distress and to make himself understood. . . only this conscious thinking *takes the form of words, which is to say signs of communication*, and this fact uncovers the origin of consciousness. In brief, the development of language and the development of consciousness. . . go hand in hand. . . . The human being inventing signs is at the same time the human being who becomes ever more keenly conscious of himself (*The Gay Science*, 1882, pp. 298-299, translated by Kaufmann, 1974).

The following fragment, found in the manuscript for *The Will to Power*, is every bit as explicit and indicative of his vision of language in relation to (rational) thought, going into greater detail about how the former not only generates the latter, but how it also confines it within a certain guiding—as well as limiting—linguistic structure:

> Now we read disharmonies and problems into things because we think *only* in the form of language—and thus believe in the "eternal truth" of "reason" (e.g., subject, attribute, etc.). *We cease to think when we refuse to do so under the constraint of language; we barely reach the doubt that sees this limitation as a limitation. Rational thought is interpretation according to a*

> *scheme that we cannot throw off* (1886-1887, p. 283, translated by Kaufmann and Hollingdale, 1968).

Polish researcher Alfred Korzybski appears alongside Nietzsche as a precursor to Whorfianism in Pula (1992). This doesn't come as a surprise, since he was mainly known for his work on "general semantics", a field of his own creation that echoes linguistic relativity's presuppositions. In 1933, he stressed the "tremendous power the structure of an habitual language has" and how it "enslaves us", finally asserting that "the structure which a language exhibits, and impresses upon us unconsciously, is *automatically projected* upon the world around us" (apud op. cit., p. 52).

More typically proposed predecessors to Whorfianism are seventeenth-century philosophers Leibniz and Locke (Koerner, 2000, p. 3), as well as eighteent-century's Diderot and Condillac (Lucy, 1997, p. 293). Yet the most cited works involve the German Romantics. German philosopher and linguist Wilhem von Humboldt (1769-1859) has been commonly credited as *the* forefather of the linguistic relativity principle (Miller, 1968, p. 10), mainly through the development of his concept of *Weltansicht*, which establishes that in each language "resides" a different *worldview* (*Weltansicht*) that guides the "sensitivity of thought and feeling" of their speakers (Underhill, 2009, p. 99). This notion is not to be confused with *Weltanschauung*[6], since they have "two clearly separate meanings": *Weltansicht* refers to "the patterning of conceptual frameworks and the organisation of ideas which makes up the form of the language ... within which we think and without which we cannot think in any conceptual or sophisticated manner"; *Weltanschauung*, however, expresses "the intellectual refinement and elaboration of those fundamental conceptual frameworks which enable us to give form to various mindsets or ideologies" (op. cit., p. 106).

2.1.1 The turning point

Albeit Humboldt is often said to be responsible for these ideas, "the image itself of language as the *organ of thought* was not coined" by him, "but inherited from [fellow German Romantics] Jean-Georges Hamann and Herder" (op. cit., p. 59). The former "had already asserted that 'the entire capacity to think rests on language'", while the latter had "the idea that 'the differences between languages parallel differences in ways of thought'"(Langham Brown apud ibidem). This last quote is *crucial* to understand our argument for setting apart linguistic relativity from Whorfianism. Very succintly, it encapsulates the basic premise of the linguistic relativity principle, and more eloquently than Whorf's own formal definition:

> We are thus introduced to a new principle of relativity, which holds that all observers are not led by the same physical evidence to the same picture of the

[6]As we can see happens in Miller (op. cit.).

> universe, unless their linguistic backgrounds are similar, or can in some way be calibrated (1940a).

What this comparison exemplifies is clear: Whorfianism as an academic tradition wasn't born out of the first straightforward formulation of the principle linguistic relativity rests on, since that already happened more than a hundred years prior to Whorf's take on it, as Herder's writing has shown us; what's more, Herder's was followed by Hamann and Humboldt's own visions, then Nietzsche's, then Unamuno's, then Korzybski's, and so on and so forth. If all of these viewpoints refer to an almost identical premise, rendering them interchangeable, then why is it that *only* Whorf's[7] functioned as the theoretical basis of a whole new line of research? The answer, we maintain, lies not in *what* was presented, but *how*.

We should consider two factors in dealing with this circumstance. First, that copies of *Science and Linguistics* (1940), from where we took the above definition of linguistic relativity penned by Benjamin Lee Whorf, have been circulating through university classes and other academic settings for many decades—many of us can recall a course or two where the teacher handed out a version of it, be it the complete article or just its famous diagrams and illustrations. Its format, similar to that of a pamphlet or a booklet (brief, pictoric, informative), lends itself to mass distribution.

Secondly, unlike any of the previous proponents of relativism, Whorf substantiates his proposition with concrete practical examples. These included the controversial characterization of Hopi speakers as having no concept of "time" as a direct consequence of their language's configuration, Eskimo's various terms for *snow*, and the discussion of the categorization of color. All of these practical implications exemplified to illustrate his theory are what impelled Brown and Lenneberg, coming from Harvard and MIT respectively, to treat it as a testable hypothesis in their 1954's *A Study in Language and Cognition*, effectively inaugurating with its publication, *ipso facto*, a new research field.

The noticeably wide gap between Whorf's publication of his article (1940) and Brown and Lenneberg's paper more than ten years later obeys to the republication of some of his writings (*Science and linguistics* included) in a posthumous collection in 1950. *Four articles on metalinguistics*, which they cite in their references, brought Whorf's work into the limelight well after his death:

> The republication of Benjamin L. Whorf's articles on what Trager calls metalinguistics has aroused a new interest in this country in the problem of the

[7]Probably influenced by the work of his teacher, Edward Sapir, who contemplated the possibility of language having a transformative effect on thought, although he ultimately relegated it to being only a "symbolic guide to culture" (1921, p. 210), not much else than "a guide to 'social reality'" (1929, p. 209); that is, a reflection of the particular customs and views of a given cultural community through symbolic (linguistic) means, which could potentially enforce "certain choices of interpretation" (op. cit., p. 210). His writings rather concentrated on arguing in favor of the relevance of linguistics, for this very reason, as a useful tool for sociological and anthropological studies (1924, 1929).

> relationship that a particular language may have to its speakers' cognitive processes (Lenneberg, 1953, p. 463).

In that same year, a manuscript for *An American Indian model of the universe*, another article discussing Hopi as a "timeless" language, probably written in late 1936, was also published (Carroll, 1956, p. 18).

2.2 Whorfianism's ill-defined object of study

Brown and Lenneberg's study set a key precedent in linguistic relativity research at the same time that they pioneered it, establishing a paradigm, or tradition, that would later be referred to as *Whorfianism*, among other names already cited here. Their 1954 paper, taking the aforementioned Eskimo's categorization of snow as an example, attempted to test whether Whorf's claim that linguistic variation parallels cognitive differences between speakers of different languages held true. Their own experiment dealt with a subject already discussed in his writings: color categorization. Interestingly enough, they start by refuting Whorf's rationale:

> What can be said of the English speaker's ability to distinguish the same three kinds of snow? When different stimuli do not ellicit differential responses, the stimuli may or may not be discriminated. A subject may be perfectly able to distinguish two situations and still not care to do anything about it. Consequently the fact that English speakers do not have different names for several kinds of snow cannot be taken to mean that they are unable to see the differences. It would seem, then, that all such comparisons are psychologically inconclusive. The Eskimo and American may or may not see the world differently (p. 455).

This is a manifest rejection of Whorf's distorted view of semantic categorization; specifically, they oppose the assumption that each language providing a default semantic arrangement of reality proves a transformative effect on the perception of their speakers. This, as Brown and Lenneberg remark, would be an illogical conclusion, since linguistic performance cannot possibly prove differences in basic perception—as they argue, linguistic expression by itself doesn't offer an insight into the individual's visual ability, for instance.

Rather, they go on to predict that "increased frequency of a perceptual categorization will mean a generally greater 'availability' of that category" (p. 456). This points to the existence of an ordinary linguistic mechanism, not to a transformative quality with regards to perception: words or expressions that refer to perception will be retrieved faster if they are frequently used. This is probably true for *any* word or phrase regardless

of what their denotational values are. However, this wasn't enough to dispel Whorf's misjudgement; instead, both his particular view of categorization and the methodology employed by Brown and Lenneberg were adopted in linguistic relativity research as a standard.

Lucy (2016) concurs with considering their experiment being as influential as we have posited, pointing out that the "methodology used in this color research remains an enduring legacy of this strand of linguistic relativity research" (p. 491), even though we contend that it has had an enduring effect on *all* of linguistic relativity research, as it inaugurated its most dominant empirical strategy: domain-centered approaches[8]. These consist in choosing a perceptual domain of experience and studying how different languages encode or "categorize" it. Striking contrasts between languages are claimed to echo a difference in the perception of reality. Whorf commented on domains such as substances (the various words for snow in Eskimo) or time (the purported lack of such a notion in Hopi), while Brown and Lenneberg focused on color, as we've seen.

Of course, this makes for an invalid strategy, for it mistakes semantic categories that rest on perceptual nuances for perception *itself*, just as we mentioned earlier. Nonetheless, it is a very easy one to adopt. The very same logical misconstruction can be exploited not only to assess substances, time or color, but *any* other semantic category or denotational value marked by perception that languages present variation in: orientation, motion, sound, textures, objects, emotion, number, etcetera.

The fact that every language adopts different conceptual models to represent reality cannot possibly prove that reality is *perceived* differently, as we argued. Otherwise, and taking Pullum's (apud Pinker, 1994, pp. 64-65) examples, one would necessarily have to maintain that horsebreeders having "various names for breeds, sizes, and ages of horses", or "botanists" having "names for leaf shapes", indicates a significantly different way of perceiving reality, instead of it just being a useful way of classifying frequently-seen objects through language.

What's more, adopting Whorf's conception of semantics would make it impossible to attribute a cause to these contrasts in categorization: in place of attributing it to being frequently exposed to such objects, one would have to defend that having several words for them both cause *and* reflect basic differences in the perceptual abilities of horsebreeders and botanists. This, of course, would condemn us to a circular reasoning that argues that they have different words for leaves and horses because of their unique perceptual qualities, and that this singular perception is caused by having different words for leafs and horses, at the same time.

[8]The terminology comes from Lucy (1997, p. 298).

2.2.1 Widespread adoption of Whorfianism

To explain the reason behind the general adoption of these practices and theoretical assumptions, we should start by pointing out that linguistic relativity has been traditionally defined by having a problematic relationship with empirical analysis. Lucy (1997) attests to the *quantitative* insufficiencies present throughout the past century:

> Surprisingly, there has been an almost complete abscence of direct empirical research through most of the present century.... The neglect of empirical work is so conspicuous that it must be regarded as one of the central characteristics of this area of research... (p. 294).

According to Everett (2013), the *qualitative* aspect of this kind of work has been historically lacking as well—a situation that has only changed in recent times: "Perhaps surprisingly, despite the deep historical roots of the question at hand, quality empirical data have only been arrived at somewhat recently (p. 1)". Lucy (2016) has stated that these issues have been generally resolved or mitigated considerably, given that, "over time, the quality, quantity, and range of empirical research on linguistic relativity have steadily increased" (p. 508).

Undeniable as it is that empirical studies have gradually increased in number, we believe that their qualitative defects persist. If one measures the quality of experiments in this field by the degree to which they can convincingly demonstrate significant linguistic effects on general cognition, the verdict seems to be clearly negative. In fact, this generalized reluctance to produce *any* kind of experimental work in the past stemmed from the very same obstacle—the abscence of an adequate and cohesive theoretical framework—that has impeded the emergence of convincing results and data, promoting heavily contentious ideas instead.

Without a standard theoretical basis to rely on, an initial timidness to conduct practical studies is to be expected. The indeterminacy of linguistic relativity's object of study has also caused it to slowly build its empirical tradition on top of the incongruencies underpinning the paradigm set by Whorf and unwillingly spread by Brown and Lenneberg. The notable rise in empirical work in the past three decades can be explained by the broad acceptance in psychology of that paradigm—a process that was probably accelerated by publications from Boroditsky (2001, 2003) and others, as well as the subsequent media coverage and cultural products[9] that they lead to:

> One key shift has arisen as most research activity has moved to psychology, where linguistic relativity is now a well-recognized, if controversial, focus of research... This disciplinary shift then affects the course of research, because

[9]Engle (2016) gives an account of the role American film *Arrival* had in the most recent resurgence in popularity of the "Sapir-Whorf hypothesis."

> most psychological studies lean heavily toward a domain-centered approach. Hence, we find most reviews anchored around domains, such as color, space, time, motion, and objects, but none anchored around language structures—which parts of speech, lexical classes, grammatical categories, or indexical forms affect thought (Lucy, *op. cit.*, p. 499).

This "shift", as Lucy calls it, is exemplified by Boroditsky's (2003) survey of recent literature at the time, which also served as a foundational guideline of sorts for linguistic relativity research in psycholinguistics, since it offered an enticing question to present its premise ("does language shape thought?") in addition to a formal classification of semantic domains (space, time, shapes and substances, objects) that had already been studied. Likewise, Wolff & Holmes' (2011), and Everett's (op., cit.) reviews center *exclusively* on domain-based approaches, including other categories such as number, gender, and emotion.

Lucy suggests that widespread insistence on sticking with domain-centered approaches, rather than focusing on structural variation between languages, is the main culprit for the disorderly state of affairs linguistic relativity research finds itself in:

> There are, of course, individual studies that do take structural differences as their point of departure.... But these studies, typically emerging among bilingualism researchers for whom structural contrasts are salient, do not form part of a systematic disciplinary effort to explore which types of categories matter and why.... This drift away from language structure hinders the theoretical integration of the research enterprise... (op. cit., pp. 499-500).

Considering that we've already proposed a systematic classification of all relevant linguistic structures as one of the pillars of our own theoretical framework, we can only agree with the importance Lucy attributes to their identification and scrutiny. But the sole decision of choosing a structure-centered approach to conduct research won't suffice on its own. The main condition that has to be complied with in order to prove lingusitic relativity, rather than the "point of departure" being the cause (linguistic structure) or the effect (perceptual domain), is that the two are causally linked by an anatomical and functional examination of the brain during a task that involves only the latter. If, instead, the experiment involves a task resolved by resorting to linguistic mechanisms, and performance is the only measurement tool, adhering to a structure-centered strategy or otherwise is absolutely irrelevant.

To illustrate the latter case, we will review an example of structure-centered research picked by Lucy (ibidem), Athanasopoulos & Bylund (2013). In this case, the chosen linguistic structure is grammatical aspect, which denotes the realization of an event or the extension of a given state in relation to time. Aspect marking is posited as a potentially

influencing factor on cognition related to motion events. The rationale behind this prediction can be summed up as follows: since Swedish differs from English in that it doesn't explicitly encode progressive aspect (the ongoing status of events or states, also refered to as *ongoingness* here) while the latter does mandate its marking (through present participles, like *writing* or *cooking*), their speakers are expected to display analogous differences in how they process ongoing motion events[10].

The experiment is based on the hypothesis that, when presented with an ongoing event, like a car or a person moving toward a village, native speakers of English will focus on the act of *driving* or *walking* (its ongoingness), while Swedish speakers will focus on the eventual arrival at the village (its endpoint). Given the markedness of English with regards to ongoingness, their speakers are expected to be skewed towards paying greater attention to the motion event itself as it is unfolding rather than its implicit completion or destination. The exact opposite result is expected of Swedish speakers.

In order to test the hypothesis, participants were shown several sets of three short videos: a target and two alternates. The target showed a scene with an "intermediate" level of goal orientation, where a possible endpoint for the motion event was visible but unreached during the clip. On the other hand, alternates were positioned at the extreme levels of goal orientation (low and high): the lower one, labelled [-endpoint], depicted the motion of objects or people along a trajectory with no visible endpoint; the most goal oriented scenes, tagged by the authors as [+endpoint], showed an entity reaching its destination (for instance, a person entering a building).

First, participants were shown the videos in a sequential order (one at a time), prompted to describe each of them verbally by asking what was happening in each scene. As expected, Swedish speakers mentioned endpoints more frequently than English speakers. Next, another set of participants underwent the same task, but this time it was designed, in principle, not to rely on linguistic processes: they had to respond from memory with similarity judgements. Likewise, Swedish speakers showed a greater tendency to match the target clip with the [+endpoint] alternate than English speakers.

Finally, an additional set of participants went through the same task, but this time there was linguistic interference while viewing the videos: participants had to listen to a string of three two-digit numbers (for example, 11, 45, and 65) and were asked to repeat them out loud until the videos in each set ended; then, they were asked to proceed with the similarity judgement. Previous differences between Swedish and English speakers didn't occur. The authors claim that these findings are "in line with a large body of literature that shows that cross-linguistic differences disappear under a verbal interference condition" (p. 300). They interpret these results as having significant implications for linguistic relativity:

[10]It is noted that Swedish can in fact convey progressive aspect, though in a circumlocutious manner (p. 292).

> . . . here we demonstrate that even in the abscence of any linguistic prompts in the instructions, cognitive differences between speakers of different languages do emerge. This is arguably because our memory task posed greater cognitive demands on the participants, thus encouraging and increasing reliance on verbal strategies to solve the task (ibidem).

Conversely, we believe that these results support our view that such a methodology is plainly inadequate for linguistic relativity research. To begin with, asking participants to rate scenes based on similarity—ranging from "very dissimilar" to "very similar" (p. 296)—constitutes giving linguistic instructions. More importantly, the fact that alleged differences in cognition fully depend on verbal strategies (and disappear when these verbal processes aren't available) indicates that it is not a contrast in general cognition what is being measured, but a difference in the strictly linguistic representation of the stimuli. This method, as a consequence, can only demonstrate differences in the default conceptual representations induced by each language:

> We surmise that language structure. . . provides the individual with a system of representation, some isomorphic version of which becomes highly available for incorporation as a default conceptual representation. Far more than developing simple habituation, use of the linguistic system, we suggest, actually forces the speaker to make computations he or she might otherwise not make (Pederson et al. apud Gleitman & Papafragou, 2012, pp. 8-9).

Our conclusion is backed by Lucy's (1997) statement that "linguistic relativity is not the same as linguistic diversity. *Without the relation to thought more generally (i.e. beyond that necessary for the act of speaking itself)*, it is merely linguistic diversity" (p. 295, emphasis mine). If the claimed cognitive effects cannot be isolated from language use, they simply don't exist. Just like domain-centered approaches, structure-centered ones are characterized by failing to evaluate anything other than purely linguistic variation. This is a direct consequence of Whorf's initial formulation of relativism, and it has lead authors like McWhorter to underestimate the scientific prospects of linguistic relativity:

> Their experiments are clever and elegant, and only the most rabid skeptic could deny that their work has shown a connection between language and thought. Yet most would consider it a fair assesment that the work of this cohort, often termed the "Neo-Whorfians," has shown that language's effect on thought is distinctly subtle and, overall, minor (2014, p. xiv).

It's not that the centuries-old debate about the connection between language and thought has finally provided an answer—and a dull one, at that; instead, this shows how this sort of research has been mistaken for linguistic relativity, when in practice it's more of a

specific branch of linguistic typology, centered on certifying the existence of structural and functional traits by empirical means. Our goal is to lay out a theoretical framework that can sharply separate the two.

3 Towards a theoretical framework for linguistic relativity research

So far, our work has taken a purely descriptive course of action. Going through the historical development of linguistic relativity as a scientific discipline has allowed us to identify the root causes of its major flaws, as well as outlining them more clearly. Now that we have diagnosed the obstacles this area of study faces and uncovered their origin, we can make the jump from their descriptive evaluation to a more constructive endeavor: the formulation of a model that provides a systematic classification of linguistic factors and the effects they have on general cognition.

Our proposal isn't entirely dissimilar to the formal organization of semantic domains (time, space, orientation, etc.) we have already reviewed. The domain-centered classifications that characterize this field's academic tradition do share a similar goal with our own model: making sense of the transformative power language exerts over our experience of reality. A point where they strongly diverge is their focus: while Whorfianism has centered around cognitive *effects* (purported differences in perception), our framework takes the linguistic *causes* (differences in language mechanisms) as its ordering parameters.

However, this isn't the main qualitative contrast between the two perspectives, but mostly a formal distinction. We could have opted for arranging our theoretical framework by expected cognitive effects rather than by their causes. As we have already established, it isn't empirically relevant whether we depart from language to arrive at cognition or viceversa, as they point to each other in this particular relationship. Nevertheless, we have to concede that, in a logical sense, it certainly seems to be a sounder option to go from cause to effect than in the opposite direction.

Crucially, the fundamental distinction that sets our model apart from other approaches in linguistic relativity is the formal definition of those causes and effects; that is, the establishment of precise conceptual boundaries around its objects of study. Neither structure-centered nor domain-centered strategies have come close to attempt such a thing. Conversely, our proposed framework rests on postulates that try to provide the theoretical integrity and definiteness that Whorfianism has notoriously lacked.

3.1 A first approximation to our framework

The elaboration of a theoretical framework for linguistic relativity must accomplish three things: (1) clarifying *what* constitutes a relativistic relationship between language and cognition, (2) presenting a workable methodological process that aligns with that conception, and (3) offering a formal classification of linguistic aspects that allows for their organization as minimal, empirically-testable units, so they can better fit experimental activity. In order to reach the first milestone, we have spent the previous sections of this paper explaining what relativism is *not*. Taking into account the arguments we have laid out, we can now briefly define what it *is*: the study of the effects of linguistic phenomena on *general* cognition.

3.1.1 Methodological process

It is necessary, as a second step, to further refine our vision of linguistic relativity by establishing the practical methodology whereby its occurrence can be unambiguously proven. Since language pervades many fundamental aspects of the human condition, the generation of hypotheses, along with the discussion of experimental results and their implications, *must* come from a multidisciplinary perspective. The contribution of linguistics, cognitive science, psychology, philosophy, anthropology, neuroscience, and any other discipline that can touch upon our objects of study in any degree are critical to experiment design and theoretical substantiation.

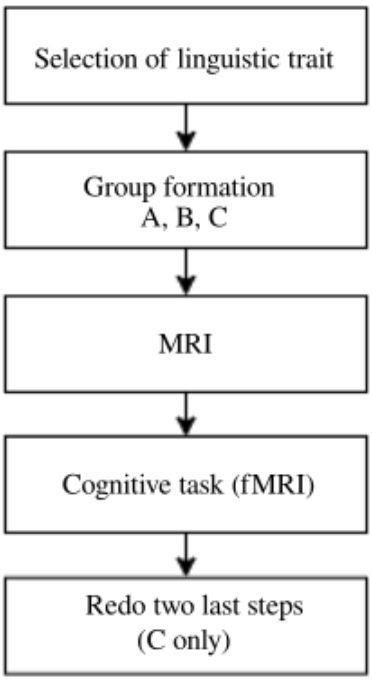
On the contrary, the empirical verification of these hypotheses has to rely on a narrower approach, involving language learning (whether L2 or Ln), and the use of neuroscientific tools. To the best of our knowledge, language learning has only been considered in this field to the extent that researchers were unsure of its adherence to the ill-defined conception of the linguistic relativity principle that we critique in this paper. Athanasopoulos (2012) ratifies our supposition when he claims that recent research in regards to language learning "has *begun* to investigate the question whether someone who learns a second language may also acquire new concepts and gain a different perspective of the world" (p. 5, emphasis mine).

Instead, we stand by the idea that *only* language learning can possibly prove direct causation in the investigation of the transformative qualities of language regarding cognition. Hence, we use language learning as a verification tool in our proposed methodological process, a graphical representation of which can be seen in Figure 1. In going through the entire procedure, we will expound its critical role.

First, a potentially transformative linguistic trait has to be selected. Next, four groups are formed based on their competence in a language that presents that specific trait. Group A is comprised of participants that are fluent (native or otherwise); group B, of individuals that aren't competent in languages sharing that characteristic; finally,

group C includes a subset of participants from group B that will later become competent through language learning. Then, an MRI is practiced on all groups in order to detect anatomical differences between them. If they exist, data extracted from this process can help in broadening the initially expected contrasts in cognition between groups, as well as facilitating the design of a cognitive task in a more precise way to account for them; if they don't, the cognitive task is designed to involve only the processes that were originally hypothesized to be influenced by the selected trait. An fMRI will map the activity on the participants' brain so as to check particular functional patterns that could differ across groups. It is crucial that these tasks involve cognitive processes that extend beyond the mere use of language.

Figure 1: A graphical representation of the methodological process.



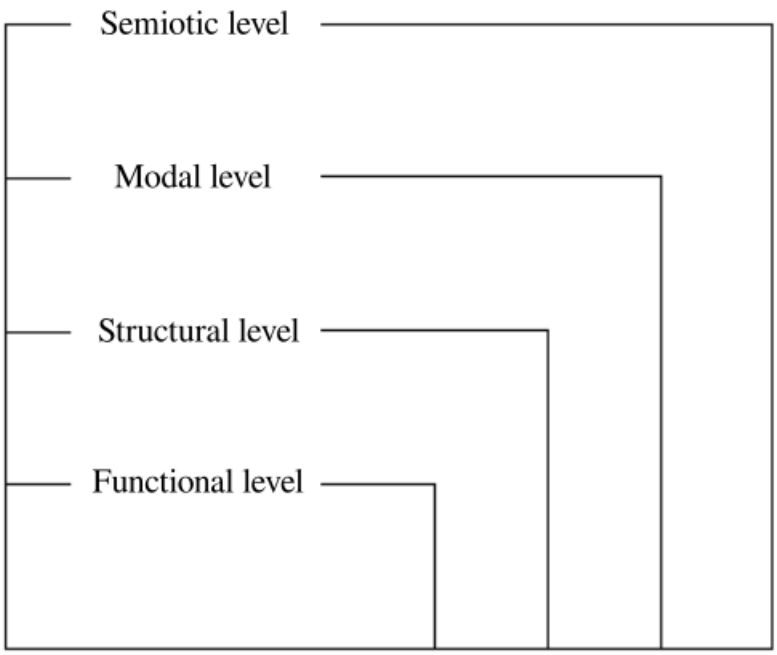
After members of group C have acquired a sufficient level of competence (fluency), they redo the two final steps. If neuroscientific tools show alterations in brain structure and function that resemble those in participants of group A, direct causation beyond correlation can be unequivocally verified. This part of the process is crucial, since working only with non-natives who have acquired the language but haven't been tested before doing so can't prove direct causation with the same degree of confidence as undergoing before-and-after examinations. An example of the validity of our approach can be found in Hölzel et al. (2011), which had sixteen participants undergo an eight-week mindfulness meditation program. MRI scans were produced before and after completing the program,

showing an increase in grey matter in different regions, such as the anterior cingulate cortex (ACC) and the hippocampus. Fox et al. (2014) found, in a meta-analysis consisting of twenty different studies, that eight brain regions were altered consistently through the process of learning to meditate.

3.1.2 A formal model of language

The third objective, building a formal classification of linguistic traits, is vital to make good use of the methodology we just described. Firstly, we have to arrange the whole extent[11] of linguistic traits in different hierarchical levels—this format will help us segregate them and illustrate their relationship with one another, like we can see in Figure 2. As we just mentioned, it is absolutely necessary to divide them into minimal testable units so we can formulate a hypothesis for every single one of them at a time, unequivocally and without overlap. This will also have the benefit of facilitating the generation of new hypotheses on a systematical basis, regardless of the language at hand. In place of scrutinizing the mechanisms of each language we may be dealing with and trying to come up with a way to isolate their potentially influencing qualities, we can use our universal model as a template.

Figure 2: A graphical representation of our language model.



[11] Linguistic relativity research has a tendency to deal exclusively with oral languages.

We will now proceed with the description of every one of these levels and which characteristics of language they are comprised of; later, we will go through each one of them individually, offering examples of published research that adjust to their criteria as well as projecting new lines of study that could arise by adopting their use.

Each level represents a different scale, in terms of complexity, in which language as a whole can be analyzed. We take most of their labels from Lucy (1997) for their convenience, though we apply considerably different rules to demarcate them. They are arranged in a hierarchical fashion, ordered from top (language in its most basic conception) to bottom (most complex). Every level past the semiotic is necessarily contained within all the previous ones, as they all presuppose, sequentially, a more basic form in which they are integrated. Thus, one could describe them as ranging from a "farther", more general view of language to a "closer", more concrete one. Like a pointillist painting, standing afar gives you the overall picture, while the closer you look into it, the more you can distinguish its constituent parts.

The first one, the *semiotic* level, is defined by Lucy (p. 292) as concerning "how speaking any *natural* language at all may influence thinking"[12]. Instead, we propose that it centers around the possibility of acquiring any language, natural *or* formal, influencing cognition. We will see some instances where learning a formal language, such as musical notation, does indeed alter the brain's morphology and functioning. This level is the most elemental, as it is based on the contrast between individuals that have gone through the acquisition of any language whatsoever—that is, of *any* symbolic system—and those who haven't. Inbetween both extremes, delayed language acquisition would also fit in this level.

The second one, the *modal* level, focuses on the mode in which an acquired language is represented (oral, written, signed, touch-based, etcetera) and the possible effects on cognition its variance could have. A paradigmatic case study would be gauging the effects of literacy (written language acquisition) or learning a sign language.

The *structural* level groups mechanisms different languages operate with, within a given mode; that is, it concerns crosslinguistic variation inside specific modes of articulation. For instance, in written languages, the contrast between representational systems (logographic, alphabetic, syllabic...); in oral languages, the presence of tonality; in both of them, declension; in every possible mode, syntax rules or semantics. This is the most productive level, for it has the greatest amount of variation.

Lastly, there is the *functional* level, which revolves around differences in usage within the same language. The use of metaphors and other rhetorical figures, speaking in a certain pitch pattern, or organizing a discourse in a certain way would serve as good examples, as they don't constitute linguistic structures by themselves, but different manners in which they can be arranged to convey a specific meaning or produce a desired

[12]Emphasis mine.

aesthetic effect.

We can see how every single linguistic mechanism past the semiotic level is encased within the previous ones by going on a regressive path. If we take metaphors as an example, their inclusion in the functional level presupposes that its occurrence, in practice, is made possible by the acquisition of a symbolic system (semiotic level), that is articulated in a specific mode (modal level), in which a particular language provides a collection of structures (structural level) that can be used in a certain way (functional level) for different purposes.

This hierarchical constriction guarantees that the facet of language we analyze isn't treated as an entirely abstract and general phenomenon, but something that has a concrete way of taking form and that can be described in terms of how it interacts with language observed from every possible angle; moreover, even though figurative use or syntax might be present in every language, they might not cause the same exact effects depending on the mode or structures they are realized on. This gives us a greater empirical integrity.

One final detail we should comment about our model before we provide concrete examples is that variation in every level except for the functional one can predict *permanent* neural alterations, as long as linguistic competence is upheld, whereas variance in the functional level can only attest for *temporal* effects, since it encompasses different uses of language, and not the implications of its acquisition, as the rest do.

3.1.2.1 Semiotic level

Between individuals, the most basic difference in language that can be relevant to linguistic relativity research is its very presence. The acquisition of a symbolic code (any cohesive set of form-meaning pairings) is bound to have a significant impact on how the brain works. In fact, Romeo et al. (2018) found that earlier linguistic exposure and participation during childhood is correlated with a greater connectivity between the left arcuate fasciculus (AF) and the left superior longitudinal fasciculus (SLF), two white matter tracts that are mainly associated with language processing and motor planning. From their results, one can gather that the opposite is also true: that the later the occurrence of regular exposure and participation, the less prominent the neural alteration will be.

However, delineating the exact changes first-language acquisition causes in the brain is a difficult enterprise, as the number of people who haven't been exposed to language is very scarce. Moreover, involving them in studies tends to be "highly problematic: isolation and childhood abuse can result in all sorts of psychological problems, which may confound conclusions drawn about lingusitic abilities" (Vishedskiy & Mahapatra, 2017, p. 15). Additionally, delayed language acquisition and consequences from brain

injuries (like aphasia) are also limited in what they can show in this regard.

Nevertheless, studies on language-deprived individuals, on rehabilitation of brain damage and delayed acquisition, as well as on earlier exposure to language, can be of utility insofar they allow for the examination of linguistic effects from the broadest, or semiotic, perspective; crucially, for linguistic relativity purposes, research must go beyond certifying the consequences these neural alterations have in language production and comprehension and design tasks that can illustrate how they modify cognition in a more general sense.

As a final note regarding this level, bilingualism and multilingualism—regardless of the specific modes and languages they reflect—could be understood as the extent of semiotic engagement in an individual. The implications of this quantitative measure could also be relevant for relativistic studies.

3.1.2.2 Modal level

The modal aspect of language refers to the various physical means through which it can be realized. Its representation can rely on signs, written or embossed characters, oral or non-oral sounds, visual cues, etcetera. This level shows great promise in linguistic relativity research, as not many studies have been published on the neural effects of acquiring different modes of articulation, probably due to the low occurrence of these learning processes among the general population.

The acquisition of literacy (the ability to read and write) in individuals that are only competent in the oral mode is a paradigmatic example of research that perfectly fits this level of inquiry. Duñabeitia et al. (2014) studied the effects of literacy on visuo-perceptual and spatial abilities, observing that they extend beyond language use (p. 11):

> Furthermore, the importance of the current set of data extends well beyond the scope of models of orthographic coding. We demonstrated that the impact of literacy on visual perception is not limited to reading (see also Szwed et al., 2012). The current results add to a growing body of evidence suggesting that literacy produces substantial anatomical and functional changes in the human brain (e.g., Carreiras et al., 2009; Dehaene et al., 2010), and that the visuo-perceptual and spatial abilities of the literate brain are not functionally comparable to those of the illiterate brain (e.g., Kolinsky et al., 2011; Reis et al., 2001; see Ardila et al., 2010, for review).

Another relevant example comes from the acquisition of tactile languages. Wong et al. (2011) show convincing evidence that a heightened sense of touch in blind individuals can be attributed to the process of learning Braille (the *tactile experience hypothesis*) to a greater degree than the abscence of vision (the *visual deprivation hypothesis*. In accordance to this, results proved that blind proficient Braille readers regularly outperformed

blind non-readers on a number of tasks involving touch, whereas the latter outperformed sighted non-readers. To further demonstrate the impact of learning a tactile language on cognition, it would be interesting to test whether sighted individuals would outperform blind non-readers *after* the former undergo acquisition, following our proposed method.

In the same vein, experiments involving the acquisition of any mode of representation can illustrate the significant influence they exert on our cognitive tendencies and abilities. Examples other than the ones discussed above include the acquisition of sign language or more unusual modes, such as the auditory (non-oral), used by Morse code and whose acquisition has been proven to alter the inferior longitudinal fasciculus (Schlaffke et al., 2017).

3.1.2.3 Structural level

The structural level reflects the most variable traits in language. Accordingly, the number of studies related to this level is the greatest among all of them. Variation in this level can come in the form of crosslinguistic contrasts regardless of mode. This includes differences in semantics, phonetics and phonology, morphology, syntax, lexicology, gestures, writing systems (phonetic, syllabic, logographic, abjad, musical notation, etcetera), and so on.

Bouhali et al. (2017) conducted an experiment whereby two groups (comprised of individuals with a high-level musical literacy and non-musicians, respectively) viewed images depicting words, faces, houses, and tools. They found that the ability to read music had an impact "on general visual processing [that] shows similarities to the changes induced by literacy acquisition", and that "such changes in perceptual abilities extend beyond music reading... [such as] improved performance with right-sided stimuli, mental rotation, and attentional processing" (p. 453). Similar findings were shown by Lee (2012).

They note, however, that albeit "unlikely, it is in principle possible that the laterality differences that we observe between groups result from differences present before any musical training, in other words that people who wind up becoming skilled musicians start off with different brains" (Bouhali et al., op. cit., p. 454). We insist on the utility of our proposed methodology to tackle this sort of problem, as it would help to determine whether learning musical notation directly causes these effects.

Another evident example of variation in the structural level revolves around different representational systems in written languages. For instance, Huang et al. (2012) compare different neural strategies involved in reading Japanese kanji characters (complex syllabograms) and Chinese hanzi ones (morphograms). Their findings, gathered through fMRIs, indicated that the greater complexity of Chinese characters were correlated with the activation of a considerably larger area of the brain while reading.

There are countless lines of inquiry within this level that are extraordinarily interesting, on top of having a great potential to paint a broad picture of linguistic relativity.

Toki Pona, a constructed language, is another such example: its grammar is deliberately minimal, with an equally deliberate semantic bias towards positive thinking. These properties could have a significant neural impact on attention, memory, emotion, etcetera. For instance, a lesser degree of polymorphism (less words, or synonyms, for similar meanings) and a simpler syntax could influence cognition; similarly, its positive bias could possibly influence neural processes and structures related to regulating emotions.

Other examples could relate to the use of complex morphosyntactic structures, like declination. Also, one could expect an increased auditory ability in speakers of tone languages (Chinese, Punjabi, Lithuanian), or a higher visual ability due to literacy in writing systems with deep orthographies—e.g., French, which has many graphemes for the same phonemes; superficial orthographies, like the one in Spanish, tend to have a single form for each sound. The acquisition of SignWriting, a pictographic writing system for sign languages, as well as Stokoe notation, which also has pictographic elements, could also constitute a relevant line of study within this level.

As a side note, this level could probably be further divided into various sublevels, but we'll leave it as is for the time being.

3.1.2.4 Functional level

Finally, the functional level refers to the practical use of language. The most salient examples are featuring the use of figurative language: rhetorical figures make use of linguistic structures in such a way that modifies their literal functions. For instance, metaphors and metonymy refer to objects of reality by using symbols in an indirect manner, not pointing toward the exact desired concepts but to others with which they share some characteristics or context. Chiappe & Chiappe (2007) indicate the tight association between metaphors and working memory, while Diaz et al. (2011) found that metaphorical sentences activate specific neural regions to a greater extent than literal sentences. Furthermore, the degree of novelty of those sentences also had an impactful effect on regions connected to memory.

These findings could be useful in the field of education. The elaboration of a didactic method that consisted in reading texts with a heavy use of figurative language and novelty could increase the capacity of readers to retain the information contained in the texts. Metaphors, pleonasms, similes, oxymorons, but also more syntax-related figures like chiasmi or anaphoras, or even rhymes, form only a small set of examples. The discoursive or formal order of a text could be relevant for this purpose as well, just like the grammatical arrangement of sentences (either more paratactical, relying on stark juxtapositions, or more hypotactical, mainly using subordination and conjunctions, in a more circumlocutious approach).

Regarding pragmatics, which also fall into this level, we'll suggest that the study of

differences in overtness may also be productive. Articulating language in a markedly indirect way, demanding the constant process of inferences, ambiguity, implicit referents and locutionary acts might result in a higher activation of neural regions and pathways. Again, these are only initial suggestions in the development of a framework that seems to be ripe with possibility.

4 Conclusion

The motivation behind writing this paper stemmed from the tension between a personal fascination with the linguistic relativity principle and a stark opposition to its academic treatment, which has failed to successfully expose the nature of a relationship as deeply reflective of the human condition as that of language and the mind. The fundamental misconceptions it relies on have constricted research to the analysis of linguistically-mediated thought, condemning the field to center around ordinary linguistic diversity instead of its truly transformative power over neural structure and function, extending far beyond the realm of language.

Our main purpose, accordingly, has been to expand the scope of the theoretical foundation of linguistic relativity as an area of study, as a means of starting to develop a complete perspective of its basic premise. In attempting to attain such an ambitious goal, we have provided a historical genealogy for the common presuppositions and practices that we have deemed conceptually unfit for that purpose, followed by their direct evaluation. Establishing their root causes has helped us outline them more efficiently while also justifying the necessity of a change of paradigm mediated by the establishment of a theoretical framework, the lack of which in this field has been previously noted by other authors.

Through the elaboration of our framework, we have tried to reach three essential milestones: first, the formal definition of linguistic relativity's objects of study; second, the introduction of an appropriate methodological process that can prove direct causation; and third, the presentation of a language model that facilitates the generation of testable hypotheses. Arranging the linguistic system in hierarchical levels has allowed us not only to find already published material that perfectly fits into each of them, but also to easily come up with new lines of study, as their particular configuration accounts for the analysis of language from all relevant perspectives: the semiotic level focuses on its presence in individuals; the modal level, on modes of representation; the structural level, on crosslinguistic variation; the functional level, on its practical uses.

Though objections and further refinements to our models are bound to take place in the future, we have made an initial effort to comply with the much-needed revision of linguistic relativity as a discipline, paying close attention to what we have identified as its essential defects in order to offer an integral solution, in hopes of opening the door to

new points of view that can help kickstart its scientific advancement.

5 Bibliography

Antzaka, A., Martin, C., Caffarra, S., Schlöffel, S., Carreiras, M., & Lallier, M. (2018). The effect of orthographic depth on letter string processing: the case of visual attention span and rapid automatized naming. *Reading and Writing*, 31(3), 583–605.

Athanasopoulos, P. (2012). Linguistic Relativity and Second Language Acquisition. In *The Encyclopedia of Applied Linguistics*. Blackwell Publishing Ltd.

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).

Athanasopoulos, P., Bylund, E., & Casasanto, D. (2016). Introduction to the Special Issue: New and Interdisciplinary Approaches to Linguistic Relativity. *Language Learning, 66*(3), 482–486.

Boroditsky, L. (2001). Does Language Shape Thought?: Mandarin and English Speakers' Conceptions of Time. *Cognitive psychology. 43.* 1-22.

Boroditsky, L. (2003). Linguisitic Relativity . *Encyclopedia of Cognitive Science, 2*(1961), 61–79.

Boroditsky, L., & Gaby, A. (2010). Remembrances of Times East: Absolute Spatial Representations of Time in an Australian Aboriginal Community. *Psychological Science, 21*(11), 1635–1639.

Bouhali, F., Mongelli, V., & Cohen, L. (2017). Musical literacy shifts asymmetries in the ventral visual cortex. *NeuroImage, 156*(March), 445–455.

Brown, R. W., & Lenneberg, E. H. (1954). A study in language and cognition. *Journal of Abnormal and Social Psychology, 49*(3), 454–462.

Caffarra, S., Martin, C. D., Lizarazu, M., Lallier, M., Zarraga, A., Molinaro, N., & Carreiras, M. (2017). Word and object recognition during reading acquisition: MEG evidence. *Developmental Cognitive Neuroscience, 24*, 21–32.

Chiappe, D. L., & Chiappe, P. (2007). The role of working memory in metaphor production and comprehension. *Journal of Memory and Language, 56*(2), 172–188.

De Cuypere, L., & Willems, K. (2008). Meaning and reference in Aristotle's concept of the linguistic sign. *Foundations of Science, 13*(3–4), 307–324.

De Unamuno, M. (1913). Del sentimiento trágico de la vida. Austral (2011).

Diaz, M. T., Barrett, K. T., & Hogstrom, L. J. (2011). The influence of sentence novelty and figurativeness on brain activity. *Neuropsychologia, 49*(3), 320–330.

Duñabeitia, J. A., Orihuela, K., & Carreiras, M. (2014). Orthographic coding in illiterates and literates. *Psychological Science, 25*(6), 1275–1280.

Elffers-van Ketel, E. (1996). The history of thought about language and thought. *Linguistics in the Netherlands, 13*(1966), 73–84.

Engle, J. (2016). Of Hopis and Heptapods: The Return of Sapir-Whorf. *ETC.: A Review of General Semantics, 73*(1), 95–99.

Everett, C. (2013). *Linguistic relativity: Evidence across languages and cognitive domains* (Vol. 25). Walter de Gruyter.

Fauvel, B., Groussard, M., Chételat, G., Fouquet, M., Landeau, B., Eustache, F., Desgranges, B., & Platel, H. (2014). Morphological brain plasticity induced by musical expertise is accompanied by modulation of functional connectivity at rest. *NeuroImage, 90*, 179–188.

Fox, K. C., Nijeboer, S., Dixon, M. L., Floman, J. L., Ellamil, M., Rumak, S. P., ... & Christoff, K. (2014). Is meditation associated with altered brain structure? A systematic review and meta-analysis of morphometric neuroimaging in meditation practitioners. *Neuroscience & Biobehavioral Reviews, 43*, 48-73.

Gleitman, L., & Papafragou, A. (2012). New Perspectives on Language and Thought. *The Oxford Handbook of Thinking and Reasoning*.

Gumperz, J. J., Levinson, S. C., Lucy, J. A., Slobbin, D. I., Kay, P., Keller, C. M., Dixon Keller, J., Bowerman, M., & Boyer, P. (1996). *Rethinking Linguistic Relativity (6 Bowerman)*.

Hölzel, B. K., Lazar, S. W., Gard, T., Schuman-Olivier, Z., Vago, D. R., & Ott, U. (2011). How does mindfulness meditation work? Proposing mechanisms of action from a conceptual and neural perspective. *Perspectives on psychological science, 6*(6), 537-559.

Huang, K., Itoh, K., Kwee, I. L., & Nakada, T. (2012). Neural strategies for reading Japanese and Chinese sentences: A cross-linguistic fMRI study of character-decoding and morphosyntax. *Neuropsychologia, 50*(11), 2598–2604.

Lucy, J. A. (1997). Linguistic Relativity. *Annual Review of Anthropology, 26*(1), 291-312.

Lucy, J. A. (2016). Recent Advances in the Study of Linguistic Relativity in Historical Context: A Critical Assessment. *Language Learning, 66*(3), 487–515.

Lee, H. Y. (2012). Exploring the association between visual perception abilities and reading of musical notation. *Perceptual and Motor Skills, 114*(3), 699–708.

Lenneberg, E. H. (1953). Cognition in ethnolinguistics. *Language, 29*(4), 463-471.

McLelland, N. (2001). Limiting the Arbitrary: Linguistic naturalism and its opposites in Plato's Cratylus and modern theories of language . By John E. Joseph . *Historiographia Linguistica, 28*(1–2), 207–214.

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

Miller, R. L. (2015). The Linguistic Relativity Principle and Humboldtian Ethnolinguistics. *The Linguistic Relativity Principle and Humboldtian Ethnolinguistics*.

Nietzsche, F. W. (1882). The Gay Science, translated by Kaufman, W. (1974).

Nietzsche, F. W., In Kaufmann, W., & Hollingdale, R. J. (1968). *The will to power*. London: Weidenfeld & Nicolson.

Pinker, S. (1994). The language instinct. William Morrow & Co.

Romeo, R. R., Segaran, J., Leonard, J. A., Robinson, S. T., West, M. R., Mackey, A. P., Yendiki, A., Rowe, M. L., & Gabrieli, J. (2018). Language Exposure Relates to Structural Neural Connectivity in Childhood. *The Journal of neuroscience : the official journal of the Society for Neuroscience, 38*(36), 7870–7877.

Sapir, E. (1921). *An introduction to the study of speech*. New York: Harcourt, Brace.

Sapir, E. (1924). The grammarian and his language. *American Mercury, 1*, 149-155.

Sapir, E. (1929). The Status of Linguistics as a Science. Language, 5(4), 207.

Schlaffke, L., Leemans, A., Schweizer, L. M., Ocklenburg, S., & Schmidt-Wilcke, T. (2017). Learning Morse code alters microstructural properties in the inferior longitudinal fasciculus: a DTI study. *Frontiers in human neuroscience, 11*, 383.

Underhill, J. W. (2009). Humboldt, worldview and language. *Humboldt, Worldview and Language*.

Vyshedskiy, A., Mahapatra, S., & Dunn, R. (2017). Linguistically deprived children: meta-analysis of published research underlines the importance of early syntactic language use for normal brain development. *Research Ideas and Outcomes, 3*(August), e20696.

Whorf, B. L. (1940). Science and linguistics. *Technology Review*, 42, 229-231.

Whorf, B. L. (1956). Language, thought and reality: Selected writings of Benjamin Lee Whorf. Edited by John B. Carroll. Cambridge, MA: MIT Press.

Wolff, P., & Holmes, K. J. (2011). Linguistic relativity. *Wiley Interdisciplinary Reviews: Cognitive Science, 2*(3), 253–265.

Wong, M., Gnanakumaran, V., & Goldreich, D. (2011). Tactile spatial acuity enhancement in blindness: Evidence for experience-dependent mechanisms. *Journal of Neuroscience, 31*(19), 7028–7037.