RIFL (2013) vol. 7 n. 3: 98-108
DOI 10.4396/20131210

Testing the untestable? Guidelines for advancing empirical research in the area of Linguistic Relativity

R. L. V. Pöhls
University College Dublin/ Universität Hamburg
victoria@poehls.com

Abstract Despite the long philosophic tradition of disputing the relation between language and thought as well as the huge amount of recent empirical research addressing it, evidence for (or against) the claim of the Sapir-Whorf-Hypothesis stays non-conclusive and the scientific community divided on the topic.
This article identifies two major methodological flaws in older and recent research related to the Sapir-Whorf-Hypothesis (or Linguistic Relativity), which can be hold responsible for this.
Firstly, the vagueness of the hypothesis itself is brought into focus and different modern readings that are precise and detailed enough to serve as testable theories for empirical research are proposed. Secondly, it is criticized that no initial agreement among researchers exists concerning theoretical guidelines determining what would count as decisive evidence for the core claim of the hypothesis. Therefore, a list of 11 guidelines for research in this area is put forward. This could, if followed, serve in refuting the pessimistic assumption put forward by researchers like De Cruz (2009) and show that the Sapir-Whorf-Hypothesis is indeed empirically testable.

Keywords: Linguistic Relativity, Sapir-Whorf-Hypothesis, Philosophic Interpretations, Empirical Testing, Test Design Guidelines

0. Introduction

The origin of the hypothesis of Linguistic Relativity (henceforth LR) – the idea that our cognition could be determined, influenced or linked to our native language(s) – «can probably be traced back to the dawn of philosophy» if one thereby only means that general «[s]peculation about the relation between language, culture and thought» (GUMPERZ, LEVINSON 1996: 1) has been entertained at least since the ancient Greeks.
While it was then a question of pure *philosophical* reasoning, Wilhelm von Humboldt's reformulation of this hypothesis at the beginning of the 19th century – «Thinking is not only dependent on language per se, but, to a certain extent, also on the specific one» (HUMBOLDT 1906: 21) – was already connected to more *empirical* considerations: while his interest was rooted in the philosophical dispute around Herder's *Abhandlung über den Ursprung der Sprache* (1772), it evolved to become core part of his comparative linguistics, thus being turned into a question

only solvable by probing the hypothesis' implications against different existing languages 'out there'.
In the 1920s, this thought was taken up by the linguist and anthropologist Edward Sapir, who stressed that he was neither interested in linguistic forms as such (e.g. how many words for «snow» there are in a particular language) nor meaning as such (i.e. can a particular language refer to something), but was intrigued to find out about the *regular* ways meaning was constructed differently in different languages and the influences this had on the worldview of its speakers. In fact, he believed that the

> worlds in which different societies live are distinct worlds, not merely the same world with different labels attached. [...] Even comparatively simple acts of perception are very much more at the mercy of social patterns then we might suppose. […] We see and hear and otherwise experience very largely as we do because the language habits of our community predispose certain choices of interpretation (SAPIR 1929: 209-210).

Sapir's idea was hereafter popularized by his disciple Benjamin Lee Whorf, who also provided the later widely criticized examples of the conceptions of time, space and matter in Hopi language, which he believed to bolster the LR hypothesis (WHORF 1956). Here, at the latest, it becomes obvious that it had by then become a commonplace that this philosophical argument around the relation of thoughts and language(s) shall be finally settled through empirical and experimental evidence rather than analytic considerations.
When Whorf's findings were criticized by linguists and empirical researchers alike Brown and Lenneberg (1953), Malotki (1983), Berlin and Kay (1969), the hypothesis soon fell out of favor. In conformity with this, the scientific community of cognitive science neglected the hypothesis in its early decades, partly because of the methodologically flawed experimental evidence, and partly because of the success of universalist and nativist research programs. Undoubtedly, the most prominent linguist figure of this line of thought is Noam Chomsky, who developed the theory of Universal Grammar (CHOMSKY *et al.* 2002).
Recently, however, the pendulum seems to be swinging back, as evidence on the effects of language differences on cognition is proclaimed to be found by different researchers in various domains (concrete domains such as color, plurality, individuation or spatial frames of reference and more abstract domains such as time, gender, integers, or emotions are only the ones under most scrutiny (DAVIDOFF *et al.* 1999; LUCY 1992a; IMAI, GENTNER 1997; LEVINSON 2003; MAJID *et al.* 2004; BORODITSKY 2001; BORODITSKY *et al.* 2003; GORDON 2004; PICA *et al.* 2004; FRANK *et al.* 2008; WIERZBICKA 1999).
Nevertheless the area stays controversial as negative evidence is also available (HECK *et al.* forthcoming; SCHUH 2011; JANUARY, KAKO 2007), and the same experimental results are interpreted to lead to opposite conclusions concerning the validity of LR (PAPAFRAGOU *et al.* 2007). While giving empirical evidence derived from specifically designed experiments is agreed upon between supporters and critics to be the method of choice to finally reveal if Sapir and Whorf have been right in their assumptions, the only conclusion a non-biased observer at this point can arrive at is: there is no sound conclusion yet.

The relatively minor progress made in resolving this question might be surprising, considering that philosophical formulations of this thesis have been circulating for a

long time and in concrete forms for more than a century, and empirical research has been and is carried out to this day by a growing number of researchers. One does not have to share de Cruz's pessimism (DE CRUZ 2009), that Linguistic Relativity is not empirically testable at all, to wonder: why is this so? And might there be a more fundamental reason for the disagreement among researchers that has to be resolved before empirical research on this topic can be fruitful?

While reasons for the discord might of course be multiple[1], I suggest that two main issues lead to the inconclusiveness of the area. Firstly, the vagueness of the hypothesis itself – which interpretation is in question exactly is often not made explicit in studies professing to test it. The hypothesis put forward in early philosophical discussions might often indeed be too general to enable us to carry out empirical research. Hence it would be the researcher's responsibility to spell out which interpretation of the hypothesis is under scrutiny, but this is often neglected in empirical research. In the first section I will therefore identify the four main strands of interpreting Linguistic Relativity or the Sapir-Whorf-Hypothesis in the current literature and put forward what I believe to be a plausible interpretation of it, worthwhile to investigate.

Secondly – and this is a theoretical and methodological flaw mostly following from the first – no agreement exists concerning theoretical guidelines determining what would count as decisive evidence for (or against) Linguistic Relativism. Naturally, this leads to disagreement how empirical results should be interpreted and if experimental setups are seen as sufficient to test the hypothesis at all. In a second step I will here therefore condense the insights gained into a set of 11 guidelines for empirical research.

1. One hypothesis – four interpretations

In philosophic discussion, different formulations of Linguistic Relativity (LR) have been circulating for more than a century. In the beginning, as illustrated in the introductory quote by von Humboldt, «Thinking is not only dependent on language per se, but, to a certain extent, also on the specific one» (HUMBOLDT 1906: 21), formulations have typically been quite vague. Von Humboldt used the term 'dependent' to describe the relationship between thought and language, but the degree of dependence is not entirely specified by him. In the following this led to speculation about the nature of dependence that von Humboldt wanted to refer to (BOLLNOW 1938). But even most modern formulations – e.g. Lucy's proposal that «diverse languages influence the thought of those who speak them» (LUCY 1992a: 263) or Gumperz and Levinson's wording that «culture, through language, affects the way we think, especially perhaps our classification of the experienced world» (1996: p.1) are on their own quite general – indeed too general to enable us to carry out empirical research testing them.

It is therefore not surprising that it became particularly clear how greatly understandings of the hypothesis differed when it was first attempted to be solved empirically by Sapir and contemporaries (Whorf 1956; Brown, Lenneberg 1953; Malotki 1983). Hereinafter it was Whorf who, fearing that the thoughts of his teacher

[1] One of them might simply be related to common practice in psychological research: the assessment of the philosophical thesis per se (whatever it is) is not present in most psychologists' agenda. Often psychologists are more concerned with finding and simply naming effects of language diversity on thought, than with refining their interpretation.

Sapir were trivialized by the scientific community, most clearly pointed to the actuality that the hypothesis was and can be understood in different ways. He felt that the «proposition that an accepted pattern of using words is often prior to certain lines of thinking and forms of behavior» was being trivialized to «a platitudinous recognition of the hypnotic power of philosophical and learned terminology on the one hand or of catchwords, slogan, and rallying cries on the other» (WHORF 1956: 134).
Wolff and Holmes (2010) and Reines and Prinz (2009) give the heretofore most comprehensive synopsis of the four biggest strands of possible interpretations of Linguistic Relativity.
The most *Radical Interpretation of LR*, neither held by many current researchers nor attributable to any of the prominent forefathers, could be stated as
(RLR) languages influence psychological processes because (all) thinking depends on natural language.
This interpretation has, so it seems, mostly served as an easy-to-ridicule proposition that opponents of the hypothesis incorrectly attributed to its advocates. There exists striking empirical counterevidence (KOSSLYN *et al.* 2006 on mental imagery, HAUSER 2000 on animal cognition, SCHALLER 1991 on adults with language deficits) as well as the theoretical counterargument that to be able to learn a first language we must somehow map its words to concepts we already possess – thus we must have acquired some concepts without natural language through other cognitive faculties (see e.g. FODOR 1975, for counterarguments see DENNETT 1981, BLACKBURN 1984).

The other end of the continuum of possible interpretations is the *Trivial Interpretation of LR:*
(TLR) Language influences thoughts and/or other mental abilities, because by using language, e.g. specific words, we draw attention to things that we might happen to neglect without it.
While the truth of this interpretation was famously confirmed by e.g. Loftus' experiments on eye-witness memory's malleability (LOFTUS 1975) and demonstrates that the specific scope of this influence can be far from trivial, Reines and Prinz (2009: 1028) rightly note that even without any scientific research no one will generally doubt that language can direct attention – if you specifically tell someone to attend to something – "Look at the sky!" – this will most definitely direct their attention.
Having that in mind, it is easy to anticipate the theoretical criticism many of the above mentioned studies were facing: if this is the only way language influences thought, effects might well be seen, but this is so obvious that no new research is needed as prove.
I do agree, but also share the belief of the Neo-Whorfians (Gumperz, Levinson 1996, Boroditsky 2001, Koerner 2000) that this trivial reading misses the core of the hypothesis: our general validated intuition that strong language effects can be seen in these cases, where language is explicitly used, should rather be seen as a starting point to ask «what then might be the cumulative effect of one type of linguistic encoding rather than another being used throughout a speaker's personal history?» (PEDERSON 2010: 1017). Thus, testing the trivial interpretation of the hypothesis (TLR) can be a starting point to settle on potentially interesting research areas by establishing that language can potentially influence the domain in question, which

can then be further explored under theoretically more ambitious readings of the hypothesis.
One of those possibly more interesting interpretations is the *Habitual Interpretation of LR*. While the TLR hypothesis implies that effects only arise when language is actually used, it makes the far less self-evident and/or outrightly plausible claim that *(HLR) languages influence psychological processes because they instil habits of thought that lead us to think in certain ways by default that we would not have thought in without learning (one particular) language.*
Evidence in favor of this interpretation must show that the influence of language on thoughts extends beyond direct language use. Hence, effects of the influence must consequently manifest themselves in non-linguistic differences of behavior or (if one believes in a framework in which the mind is organized in different faculties) have a reorganizing effect outside the language faculty, for which differences in behavior would probably also be the best indication (see LUCY 1992a, KOERNER 2000).
Thus, HLR is not unattached or opposed to the claim of the TLR hypothesis, but is pushing it further: if your language requires you to describe something linguistically you will habitually and continually draw attention to it, it claims – even in situations in which language is not explicitly involved. This is then expected to have an impact on, for example, the efficiency with which one set of relations is processed relative to another or on the ease with which they are remembered.
But this interpretation becomes problematic for the researcher as

> [T]he difficulty of appraising such a far-reaching influence is great because of its background character, because of the difficulty of standing aside from our own language, which is a habit and a cultural non est disputandum (WHORF 1956: 137-138).

The most common and also quite hard-hitting criticism on evidence for the HLR interpretation is connected to this difficulty: even if some studies show diverse effects in different language groups regarding the given task, they subsequently also have to establish that the reason for this is in fact only the *language difference* and cannot be accounted for by other explanations. A comprehensive discussion of studies criticized in this way can be found in Papafragou *et al.* (2007).
Another challenge for researchers investigating HLR lies in the fact that it is only testable through comparable research between two or more language groups while TLR can be measured on a within-language group level comparing behavior when being given or denied additional semantic information.

There is one further interpretation of the hypothesis that could follow from the effects of HLR (if it proved to be correct): in leading us to habitually group certain particulars together, language could shape category boundaries and thus influence our understanding of what kind of things exist. The hypothesis to be tested would then be the following *Ontological Interpretation of LR:*
(OLR) Languages influence psychological processes because they lead us to organize the world into categories that differ from those we would discover without language.
Even though this is in general an interesting interpretation of the hypothesis (as is HLR), it is in my opinion a secondary question that should be addressed in research if HLR turns out to be true.

Thus, I conclude that HLR is an interpretation worth investigating, and that OLR would be – in case HLR had been established. On the contrary, I do not believe that the interpretation of RLR is worth studying. Although I do believe that the interpretation of TLR may lead to interesting findings in specific subfields, and despite its utility to single out possibly interesting domains for further research, I do not see it as relevant to the core question of LR.

2. Guidelines for empirical research

With more clarity concerning the possible interpretations of the hypothesis and their consequences at hand, it is possible to derive at a set of guidelines that defines what would count as decisive empirical evidence for a worthy of note (HLR or OLR) interpretation – as it should not be treated as a matter of weltanschauung, but a matter of empirical evidence if the phenomenon of LR exists.

Deriving at unbiased guidelines acceptable by all empirical researchers – whichever intuition they may initially have – thus rebuts the assertion that LR is not empirically testable at all, because no longer «[i]t remains unclear [...] what counts as decisive evidence in favor of linguistic determinism» (DE CRUZ 2009: 2) . Contrarily, I argue that if the following eleven guidelines are taken seriously, fruitful research in the field of LR is possible:

1. *LR should not be interpreted as a hypothesis about what we can think, but about what we do think.*
   Even none of the forefathers believed that we are generally limited in our thought potential, as learning a new language is always seen as a possibility to gain a new perspective. Thus, to read the claim of LR in such a radical way would clearly mean misinterpreting it.

2. *LR should not be interpreted as a hypothesis about language as an only attention-directing phenomenon.*
   This was found to be trivially true in many cases (see LOFTUS 1975). Interesting empirical research should thus focus on more than evaluating the TLR interpretation. But it can be necessary for empirical research to establish that language in any form (explicit or implicit) would have an effect on the cognitive faculty/the domain under scrutiny at all.

3. *Empirical research does not have to show that all thoughts are dependent on language.*
   Thus, only showing that the radical interpretation is wrong cannot refute LR as a whole. As mentioned above, mental imagery such as the mental rotation of 3D objects (KOSSLYN *et al.* 2006) is widely believed to not rely on language at all. Other possible candidate domains include visuo-spatial reasoning in mapping tasks or creative processes like composing music (SACHS 2007).

4. *Research that finds the influence in one particular domain (e.g. color, numbers, space) is evidence for LR confined to that domain.*
   and

5. *Research that finds the influence on one particular difference of behavior (e.g. memory, spatial orientation) is evidence for LR with respect to that kind of behavior.*
   Claim 4 and 5 relate to a similar possible line of criticism as mentioned in 2: every critic believing otherwise does confine himself to the too rigid RLR interpretation – and rules out the likely possibility that LR could have an influence on some domains and behaviors, but not others or that the strength of influence could vary across those. On the other hand, these claims should also prevent overstating: if an effect is found in one domain influencing one particular behavior, this only shows that LR has an effect in the tested condition, but cannot be universalized.

6. *The hypothesis cannot be verified due to the observation that very culturally or environmentally specific objects (e.g. historically situated artifacts or culture specific natural kind categories) are received in varying manners. So, LR must apply to cognitive domains that operate on 'basic' and 'common' experience.*
   It should not come as a big surprise that the conceptualization of a culturally created object/ an environmentally specific class of objects is different in a culture/environment where it is unknown. I do understand that the notion of 'basic' and 'common' is far from unproblematic. However, I think it is possible to agree on domains that seem reasonably basic to cognition and human experience, but are «just shy of exhibiting a universal pattern» (PEDERSON 2010: 5). Although one may also turn out to be wrong about the hypothesized basicality, it is possible to discriminate promising candidates (color, space, or in my experiment: the human body). Choosing such a domain helps to keep other variables such as varying attentional engagement or dissimilar task satisfying abilities in the groups under investigation to a minimum.

7. *As LR is a hypothesis essentially concerned with differences,* comparative *research between two or more language groups is necessary.*
   While this has only come into question very recently in a study of Stapel and Semin (2007), it is crucial to realize that universalist's claims can only be refuted if differences between groups can be found. The erroneous conclusion that has been made in studies like Stapel and Semin (2007) is based on the fundamental assumption that the hypothesis of LR is testable using participants from the same linguistic community. With their research they wish to answer the question if «categories have a generic attention-channelling effect that affects the perception of a stimulus environment?» (*Ivi:* 25). This in itself is clearly only testing the TLR not the HLR (which can indeed be investigated by looking only at one language).
   But in their concluding section Stapel and Semin infer from their results that (1) generic features of a language drive attention to different features of a stimulus environment, and the platitude that (2) differences across languages do exist, that (3) their research provides «reliable, empirical evidence for the core of Whorf's (1975) linguistic relativity hypothesis» (*Ivi*: 31), thus the HLR interpretation.
   I believe this is not the case: they fail to show that language has any other effect apart from immediate direction of attention and their deduction is

fallacious – it is far from necessary that from (1) and (2), (3) and thus a validation of the HLR hypothesis follows. Agreeing with Lucy (1992a, 1992b) I therefore think that comparative research is necessary to make any meaningful assertions about the HLR hypothesis.

8. *Research has to make sure that languages evaluated not only give different names – in the sense of different strings of phonemes – to the same thing, but attribute different senses to their expressions. Thus, careful linguistic evaluation of competent speakers of the languages to be compared is absolutely necessary before carrying out research.*
   Although this seems nearly too obvious to be stated, this has been one of the most common flaws (see LUCY 1992b for a similar opinion).

9. *Clear evidence and a methodology of how it was obtained has to be presented on*
   a) *the habitual effect on other cognitive faculties, thus how it affects the way we interact with the world*
      and/or
   b) *the effect on our ontological representation of the world, thus how it affects our beliefs of what there is.*
      This follows from the conclusion drawn in section 3, finding that HLR and OLR are theoretically interesting hypotheses to test. a) is, as I argued, presumably easier to test, and it could therefore be the best starting point for research – but this is only a practical concern as both kinds of evidence would be valid evidence for some influence of LR.

10. *The cognitive process suspected to be affected by LR has to be assessed by* non-*linguistic measurements to avoid circularity.*
    This might be best conveyed through an example: suppose it is hypothesized that, compared to participants with a richer vocabulary, less distinct vocabulary in a specific area habitually leads to less specified thoughts and thus to a poorer recollection (even if no explicit language use occurred in the testing situation). If the recollection is then measured by a linguistic questionnaire that the participants answer in their native tongue, it is obvious that they would use less specifying terms – because these are the only available in their language! But this is far from entailing the conclusion that it had any influence on their recollection: a non-linguistic measurement (e.g. through a visual marking sheet) might have shown that their recollection was just as exact as that of the other group, but they were just not making these differences apparent through their language.

11. *Causality between language differences and observed differences of behavior has to be established.*
    This means that the difference in language must at least be the best possible explanation for the findings. The researcher has to make sure that other factors can be mostly ruled out: sought-after testing procedures should therefore make sure that in solving the task participants have to rely on (implicitly evoking) language rather than using other cognitive facilities instead. Possible competing explanations should be mentioned and it has to be shown that they cannot account for the data as well as the LR rival.

3. Conclusion

While the newly awakening interest in the Sapir-Whorf-Hypothesis and the growing body of empirical research is surely to be welcomed, it is now necessary to take a step back to evaluate the methodology behind it: without paying more attention on spelling out the exact hypothesis to be tested and instead only addressing a vague notion of the Sapir-Whorf-Hypothesis – which can and is interpreted rather differently by different scholars – it is not possible to integrate the different findings into one conclusive answer if (and if so, in which ways) languages influence thought. Similarly, unbiased guidelines determining what can count as decisive evidence, such as these presented in the last section, are a necessary second step. Only if preliminary consensus is reached among empirical researchers, whichever initial intuition they may have on the topic, can empirical results act as a valuable source of information, instead of only providing ammunition for academic turf wars.

Only with keeping that in mind, I believe that future empirical research in the area of the Sapir-Whorf-Hypothesis can be fruitful and enlightening in answering the core question of Linguistic Relativity: can our language habitually lead us to think differently and to have different ontological beliefs - thoughts and beliefs we would not have had without learning one particular language?

Bibliography

BERLIN, B., KAY, P. (1969), *Basic Color Terms,* Univ. of California Press, Berkeley.

BLACKBURN, S. (1984), *Spreading the word*, Oxford Univ. Press, Oxford.

BOLLNOW, O. F. (1938), «Wilhelm von Humboldts Sprachphilosophie», in *Zeitschrift für Deutsche Bildung*, 14 (3), pp. 102-112.

BORODITSKY, L. (2001), «Does language shape thought? Mandarin and English speakers' conceptions of time», in *Cognitive Psychology,* n. 43, pp. 1–22.

BORODITSKY, L., SCHMIDT, L. A., PHILLIPS, W. (2003), *Sex, syntax, and semantics,* in GENTNER, D., GOLDIN-MEDOW, S. (Eds.), *Language in Mind: Advances in the Study of Language and Thought*, MIT Press, Cambridge, MA, pp. 61–79.

BROWN, R. W., LENNEBERG, E. H. (1953), «A study in language and cognition», in *The Journal of Abnormal and Social Psychology*, Vol. 49, n. 3, pp. 454-462.

CHOMSKY, N., HAUSER, M. D., FITCH, W. T. (2002), «The Faculty of Language: What Is It, Who Has It, and How Did It Evolve?», in *Science*, n.298, pp. 1569-1579.

DE CRUZ, H. (2009), «Is linguistic determinism an empirically testable hypothesis?», in *Logique et Analyse*, n.208, pp. 327-341.

DAVIDOFF, J., DAVIES, I., ROBERSON, D (1999), «Colour categories in a stone-age tribe», in *Nature*, n. 398, pp. 203–204.

DENNETT, D. (1981), *Brainstorms: Philosophical Essays on Mind and Psychology*, MIT Press, Cambridge (MA).

FODOR, J. (1975), *The Language of Thought*, Harvard Univ. Press., Cambridge, MA.

FRANK, M. C., EVERETT, D. L., FEDORENKO, E., GIBSON, E. (2008), «Number as a cognitive technology: Evidence from Pirahã language and cognition», in *Cognition*, n.108, pp. 819–824.

GORDON, P. (2004), «Numerical cognition without words: Evidence from Amazonia», in *Science,* n. 306, pp. 496–499.

GUMPERZ, J. J., LEVINSON, S. C. (1996), Eds., *Rethinking linguistic relativity*, Cambridge Univ. Press, Cambridge, UK.

HAUSER, M. D. (2000), *Wild Minds: What Animals Really Think,* Henry Holt, New York.

HECK, S., KONTNY, L., LAUER, K., WITTKE, P., PÖHLS, V. (forthcoming), *Grenzen unserer Sprache -Grenzen unserer Welt?*, Studienkolleg Tübingen.

HUMBOLDT, W. VON (1906), *Humboldts Gesammelte Schriften*, Königlich Preußische Akademie der Wissenschaften, Berlin, IV.

IMAI, M., GENTNER, D. (1997), «A crosslinguistic study of early word meaning: universal ontology and linguistic influence», in *Cognition,* n. 62, pp 169–200.

JANUARY, D., KAKO, E. (2007), «Re-evaluating evidence for linguistic relativity: Reply to Boroditsky (2001)», in *Cognition*, n. 104, pp. 417–426.

KOERNER, E. F. K. (2000), *Towards a full pedigree of the Sapir-Whorf hypothesis. From Locke to Lucy*, in PÜTZ, M., MARJOJLIN, V., eds., *Explorations in linguistic relativity,* Amsterdam/Philadelphia, John Benjamins, pp. 1-24.

KOSSLYN, S. M., W. L. THOMPSON, GANIS, G. (2006), *The Case for Mental Imagery*, Oxford Univ. Press., Oxford.

LEVINSON, S. C. (2003), *Space in language and cognition: Explorations in cognitive diversity*, Cambridge Univ. Press, Cambridge, UK.

LOFTUS, E. F. (1975), «Leading questions and the eyewitness report», in *Cognitive Psychology,* n. 7, pp. 560-572.

LUCY, J. A. (1992a), *Language diversity and thought: A reformulation of the linguistic relativity hypothesis*, Cambridge Univ. Press, Cambridge, UK.

LUCY, J. A. (1992b), *Grammatical categories and cognition*, Cambridge Univ. Press, Cambridge, UK.

MAJID, A., BOWERMAN, M., KITA, S., HAUN, D., LEVINSON, S. C. (2004), «Can language restructure cognition? The case for space», in *Trends in Cognitive Sciences,* n. 8.3, pp.108–114.

MALOTKI, E. (1983), «Hopi Time: A Linguistic Analysis of the Temporal Concepts in the Hopi Language», in *Trends in Linguistics. Studies and Monographs*, vol. 20, ed. by Werner Winter, Mouton Publishers, Berlin / New York / Amsterdam.

PAPAFRAGOU, A., LI, P., CHOI, Y., HAN, C-H. (2007), «Evidentiality in language and cognition», in *Cognition*, n. 103, pp. 253–299.

PEDERSON, E. (2010), *Cognitive Linguistics and Linguistic Relativity*, in GEERAERTS, D., CUYCKENS, H., eds., *The Oxford Handbook of Cognitive Linguistics*, Oxford University Press, Oxford, p. 1012-1044.

PICA, P., LEMER,C. IZARD, V., DEHAENE, S. (2004), «Exact and approximate arithmetic in an Amazonian indigene group», in *Science*, n. 306, pp. 499–503.

REINES, M.F., PRINZ, J. (2009), «Reviving Whorf: The return of linguistic relativity», in *Philosophy Compass,* vol. 4, n. 6, pp. 1022–1032.

SACHS, O. (2007), *Musicophilia: Tales of Music and the Brain*, Picador, London, pp. 30-40.

SAPIR, E. (1929), *The Status of Linguistics as a Science*, in SAPIR E. (1958), *Culture, Language and Personality* (ed. D. G. Mandelbaum), University of California Press, Berkeley, CA, pp. 209-212,.

SCHALLER, S. A. (1991), *Man Without Words*, Summit Books, New York.

SCHUH, R. G. (2011), «Reply to Lera Boroditsky 'How Language Shapes Thought'», in *Scientific American*, n. 1. Available online: http://www.linguistics.ucla.edu/people/schuh/lx001/Discussion/d01_response_to%20Boroditsky.pdf

STAPEL, D. A., SEMIN, G. (2007), «The magic spell of language: Linguistic categories and their perceptual consequences», in *Journal of Personality and Social Psychology*, vol. 93, n.1, pp. 23-33.

WHORF, B. L. (1956), *Language, thought, and reality: Selected writings of Benjamin Lee Whorf* (Ed. J. B. Carroll), MIT Press, Cambridge, MA.

WIERZBICKA, A. (1999), *Emotions across languages and cultures: Diversity and universals*, Cambridge Univ. Press, Cambridge, UK.

WOLFF, HOLMES (2010), «Linguistic Relativity», in *Wiley Interdisciplinary Reviews: Cognitive Science*, vol. 2 (3), May/June 2011, pp. 253-265.