

Jolanta Latkowska

University of Silesia

How Relevant is the Sapir-Whorf Hypothesis to Contemporary Psycholinguistic Research?

Abstract: The paper raises the question of whether the linguistic relativity proposal, also known as the Sapir-Whorf Hypothesis, should be used as a frame of reference for modern research into the relationship of language to cognition. The question is discussed in the context of Whorf's (1956) writings, with emphasis on factors that are crucial to the proposal, i.e. language, thought, and behavior. The second issue addressed by the paper is whether linguistic categories provide an accurate window on cognition, as was suggested by Whorf and in some of the more recent debates. The analysis takes the form of a correlational study which examines the categorization criteria applied in tests that require language-based and language-neutral judgments.

Key words: linguistic relativity, Sapir-Whorf Hypothesis, cognition, language, correlational study

Introduction

To the psycholinguist and second language acquisition researcher, the *Sapir-Whorf Hypothesis*, also known as the *linguistic relativity principle*, indicates the possibility of a causal relationship between language and thought. Originally, the relationship was assumed to be deterministic, with language acting as a mould of cognition. Accordingly, the categories of the native language were attributed with formative powers which determined the perception and apprehension of extra-linguistic reality (Carroll, 1956). Because of a lack of unequivocal empirical evidence and a barrage of criticism on purely theoretical grounds (Kramersch, 2004), this 'strong' deterministic version of the hypothesis was rejected as untenable. The 'weak' relativistic hypothesis is generally accepted, however, since it resonates with a popular perception of language as a shaping but not restrictive force (Pederson, 2007). Also, the weak hypothesis received support

from a large number of studies conducted in the 1990s and during the first two decades of this century. For an informative review of relevant research see Pederson (2007) and Cook and Bassetti (2011).

Research into linguistic relativity embraced Slobin's *thinking for speaking hypothesis* (Slobin, 1996; 2003; 2004; 2005), which stipulates that verbalization induces a language-dependent mode of thinking that transforms non-linguistic conceptual content into a 'verbalizable' propositional format prior to articulation. This is done in accordance with the available linguistic categories, which additionally serve as a mechanism directing the speaker's attention to those aspects of experience that they encode. Consequently, inferences about conceptual processing may be drawn from both verbal and non-verbal performance. As research into the paradigm produced strong confirming evidence, the thinking for speaking hypothesis tends to be invoked as a *weak form of linguistic relativity* (Han & Cadierno, 2010). This reserves the term *strong hypothesis* for those strands of relativistic research that explore correlations between linguistic categories and cognitive behavior on the basis of behavioral measures that either by-pass or entirely exclude language (Levinson, 2003; Lucy, 2004).

The first decade of the 21st century saw the emergence of a bilingual turn in relativistic thought and a spate of studies investigating cross-language interaction and the ensuing restructuring of cognitive mechanisms (Athanasopoulos, 2011b; Bylund, 2011; Pavlenko & Malt, 2011; von Stutterheim, 2003). It also bore witness to renewed criticism of the Sapir-Whorf Hypothesis, as well as accusations that it was couched in terms that did not reflect Whorf's original concerns. A case in point is Pavlenko's (2014) most recent contention that Whorfian effects manifest themselves through language use and are most easily observed in bilinguals.

Such a disparity of opinion over what constitutes Whorfian effects is not new to relativistic debates. For example, research into the conceptual basis of emotion words shows that when used as stimuli in research, emotion words influence the categorization, memory, and perception of emotion. In the absence of such prompts people are unable to identify emotion and perceive it in a categorical way (Gendron et al., 2012). According to Malt and Ameel (2011), such a lack of discriminating power in a non-verbal condition is indicative of an absence of relativistic effects. Gendron et al. (2012), by contrast, disregard the non-verbal dimension and classify as relativistic the effects evoked by the presence of emotion words. *Thinking for speaking* appears to be equally contentious. While Slobin (1996; 2003; 2004; 2005) and Han and Cadierno (2010) see language-driven thinking as a (weak) form of linguistic relativity, Athanasopoulos (2011a) regards it as being solely linguistic.

This paper addresses the confusion that beclouds the notion of Whorfian/relativistic effects on cognition by referring to the writings of Benjamin Lee Whorf, who is credited with developing the hypothesis, and analyzing them in

terms of themes that are central to relativity, i.e. thought, language, and behavior. It is hoped that the analysis will help put the Sapir-Whorf Hypothesis in perspective and evaluate its tenets against current trends in psycholinguistic research. The second objective of the paper is to reexamine the extent to which language-based tests provide a neutral and/or accurate view of conceptual processes such as non-verbal categorization. This is vital for understanding the controversy surrounding the hypothesis and the rationale behind some of its subsequent modifications.

Whorf's Views on Key Relativistic Variables

Thought and Thinking. Following the recommendations of Watson, early 20th-century American psychology abandoned investigations of conscious thought and mental activity on the grounds that they were unobservable. Accordingly, a theory claiming that thinking takes place in a language and that 'the greatest light upon it [...] is thrown by the study of language' (Whorf, 1956, p. 252) must have aroused intense interest and attention, not to mention controversy and criticism. In such a climate, Whorf, who had no background in psychology and gained recognition in academia primarily as an expert on Maya hieroglyphs and Indian languages, advanced the view that linguistic diversity sparked off cognitive differences in speakers of different languages as a function of continued use of the patterns of their native language for expression. The evidence quoted in support of these claims was linguistic and behavioral, and is the subject of the following sections.

Although many contemporary analysts consider the notion of *habitual thought* to be the cornerstone of Whorf's hypothesis, it did not receive much attention in his writing. In fact, his collected works (Carroll, 1956) contain only one article devoted to the subject with only one specific explanation of what the term meant to its author. It reads as follows:

By 'habitual thought' and 'thought world' I mean more than simply language, i.e. than the linguistic patterns themselves. I include all the analogical and suggestive value of the patterns [...], and all the give-and-take between language and the culture as a whole, wherein is a vast amount that is not linguistic yet shows the shaping influence of language. In brief, this 'thought world' is the microcosm that each man carries about within himself, by which he measures and understands what he can of the macrocosm. (Whorf, 1956, p. 147)

Beautifully phrased, the quotation indicates that Whorf saw thought as the linguistic and non-linguistic content of the mind. Surprisingly, there is no mention of either regularity or frequency of occurrence that could account for the habitual nature of thought. The *shaping influence* of language is described elsewhere as a non-conscious structuring process that is not always overtly linguistic:

Thinking in a language does not necessarily have to use words. [...] Much thinking never brings in words at all, but manipulates whole paradigms, word-classes [...] ‘behind’ or ‘above’ the focus of personal consciousness. (Whorf, 1956, p. 252, footnote 1)

The now classic statement “The linguistic system (in other words, the grammar) of each language [...] is itself the shaper of ideas” (Whorf, 1956, p. 212) explains the significance of the *paradigms* and *word-classes*, which are attributed with formative powers. Whorf (1956, p. 258) stated explicitly that sentences were more essential than words.

Unfortunately, to the detriment of his theory, Whorf was often inconsistent and imprecise. In his rebuttal of the Watsonian notion that silent thinking is essentially “suppressed talking accompanied by laryngeal agitations” (Whorf, 1956, pp. 66–68), he ambiguously suggests the possibility that it is

RAPPORT between words, which enables them to work together at all to any semantic result [...] that constitutes the real essence of thought insofar as it is linguistic. (Whorf, 1956, pp. 66–68)

and that “thinking itself is in a language—in English, in Sanskrit, in Chinese” (Whorf, 1956, p. 252). In fact, vague references to the “linguistically-determined thought world” (Whorf, 1956, p. 154) abound in his texts, implicating language in general as a potent structuring agent.

This lack of precision should take no one by surprise. Whorf was clearly an open-minded researcher who was familiar with trends and developments in psychology. For instance, the idea that thought was largely linguistic was borrowed from Jung, while in *The Yale Report* (Whorf & Trager, 1996) *Gestalt theory* served as a backdrop for a discussion of the interaction between language and perception. Whorf was also aware but dismissive of alternative views of the language-thought interface. The quotations below lay out his thoughts on this matter:

Talking, or the use of language, is supposed only to ‘express’ what is essentially already formulated nonlinguistically. Formulation is an independent process, called thought or thinking, and is [...] indifferent to the nature of particular languages. (Whorf, 1956, p. 207)

Formulation of ideas is not an independent process, [...] but is part of a particular grammar, and differs, from slightly to greatly, between different grammars. (Whorf, 1956, pp. 212–213)

Nevertheless, it must be stressed that his approach to the issue was that of an early 20th-century linguist preoccupied with analyses of phonemes and morphemes and relatively uninformed about the complexity of cognitive processing which poses a challenge to interdisciplinary research a century later. Whorf (1956, p. 42) seemed to distrust psychological terminology which he regarded as a legacy of old laboratory experiments. Consequently, his use of psychology-related terms is intuitive and commonsensical, as demonstrated by the following description of how grammar influenced what a psychologist could have called perception and categorization. The specifics of both processes remain a matter of conjecture.

Users of markedly different grammars are pointed by their grammars towards different types of observations and different evaluations of externally similar acts of observation. (Whorf, 1956, p. 221)

Whorf's attempts to embrace the cognitive and linguistic levels attracted the attention of Lakoff (1987, p. 330), who called him a pioneer in (cognitive) linguistics and "the most interesting linguist of his day." Yet, from the perspective of psychology, the research did not meet the requirements of good science because it used imprecise and intuitive terminology, and employed anecdotal and impressionistic evidence, and for an obvious failure to define terms (variables) carefully and form testable hypotheses, which was vital for obtaining empirical evidence for the theory. What tends to be overlooked is that, as a linguist, Whorf was neither interested nor qualified to conduct research that met those requirements.

Behavior. The way Whorf understood the behavioral aspect of linguistic relativity had little in common with modern interpretations of the hypothesis. Under the influence of Sapir, a leading anthropologist of the time, he came to regard language as a social construct and consequently looked for signs of its influence in culturally conditioned personal and social activities. The influence manifested itself as "constant ways of arranging data" and "most ordinary everyday analysis of phenomena" (Whorf, 1956, p. 135) in line with the language habits of the community. Whorf became acutely aware of the dependency of behavior on linguistic descriptions of events in his work as a fire inspector. Although initially concerned with purely physical conditions and circumstances surrounding the starting of fires, he soon discovered that part of the blame lay with how the situation was represented linguistically. A case in point is

the now classic example of gasoline drums which, when full, require handling with caution and care. This caution is abandoned around empty gasoline drums because the word *empty* suggests a lack of danger. The sad truth is that empty drums are potentially even more dangerous because they contain explosive vapor (Whorf, 1956, p. 135). The principle behind this dependency is as follows: “people act about situations in ways which are like the ways they talk about them” (Whorf, 1956, p. 148).

The example makes clear that Whorf did not distinguish precisely between verbal and non-verbal behavior. Nor did he try to infer cognitive patterns from the observed behavior. It seems he was happy to accept that activities that were not explicitly linguistic were non-linguistic by default and that patterns of behavior reflected patterns of thought. His contemporaries, including his pupil and friend John B. Carroll, adopted a more principled approach, however, and argued for a strict separation of linguistic and non-linguistic processes. This protected linguistic relativity from becoming circular and tautological, as explained by Casasanto:

Inferring cognitive differences solely from linguistic differences is hopelessly circular. Patterns in language can serve as a source of hypotheses about cognitive differences between members of different language communities, but some sort of extra-linguistic data are needed to test these hypotheses: Otherwise, the only evidence that people who talk differently also think differently is that they talk differently! (2008, p. 67)

It also solved the problem of a linguistic bias in data analysis as researchers were able to assess cognitive phenomena in a language-neutral way, i.e. without privileging a vision of reality invoked by the language of the analysis, or indeed, the researcher’s native language. This was a problem Whorf was intensely aware of.

Language. A second dimension of the linguistic relativity principle represents the notion that languages are relative and vary in how they conceptualize and represent extra-linguistic reality. In the words of Whorf: “each language performs [...] artificial chopping up of the continuous spread and flow of existence in a different way” (1956, p. 253) and

segmentation of nature is an aspect of grammar [...]. Languages differ not only in how they build their sentences but also in how they break down nature to secure the elements to put in those sentences (1956, p. 240).

On the linguistic level, the result of the segmentation process is a large scale pattern-system of grammatical categories such as nouns, verbs, number, gender,

voices, tenses, and the like. The categories are attempts to segment and interpret experience. The rules of “patternment” and patterns of sentence structure are specific to each language and guide mental activity. Whorf borrowed the idea that the sentence is the main unit of linguistic structure and therefore should be used as a template in linguistic analyses from Bloomfield. What is more, in a truly structuralist style, he made a strenuous effort to describe carefully the linguistic terms used in his research and analyses.

Surprisingly to some, Whorf’s understanding of lexical semantics was quite vague. In fact, he did not seem to be familiar with the *Fregean* distinction between *sense and reference* (cf. Goddard, 2003). This is demonstrated in no uncertain terms by the quotation below:

That part of meaning which is words, and which we may call “reference,” is only relatively fixed. Reference of words is at the mercy of the sentences and grammatical patterns in which they occur. (Whorf, 1956, p. 259)

To put things in perspective, it should be noted that in his lifetime American linguistics was preoccupied with phonology and morphology, while semantics was underdeveloped since, like all things cognitive, it was considered to be beyond the reach of scientific investigation.

Yet another theme running through Whorf’s work was that of language as a cultural phenomenon, or more specifically, as an “especially cohesive aggregate of cultural phenomena” (1956, p. 65). It is a shared knowledge of the aggregate that makes it possible for the members of a particular speech community to communicate, as explained in the excerpt below:

We cut nature up, organize it into concepts, and ascribe significances as we do, largely because we are parties to an agreement to organize it in this way—an agreement that holds throughout our speech community and is codified in the patterns of our language. The agreement is, of course, an implicit and unstated one, BUT ITS TERMS ARE ABSOLUTELY OBLIGATORY (Whorf, 1956, pp. 213–214).

The belief that cultural norms are codified in ethnic languages inspired Whorf (1956, pp. 138–139) to address the question of an interplay of culture and linguistic norms. He was particularly interested to find out if there were “traceable affinities between (a) cultural and behavioral norms and (b) large-scale linguistic patterns.” The answer was affirmative, although Whorf conceded that the two were related rather than correlated, making the relation much weaker than some of his comments seemed to suggest. The conclusion to be drawn was that language and culture should be studied as a whole since they

were closely integrated. It is more than likely that this anthropological slant in Whorf's reasoning prevented him from using quantifiable research paradigms.

Finally, as a linguist and anthropologist, Whorf expressed concern about what Lucy (2011, p. 46) called "unwitting lingua-centrism" of linguistic evaluation and description. This is a tendency to analyze languages in terms of categories rooted in the researcher's own language. Although unable to eliminate the bias of using English as a tool of analysis in his own research, Whorf made his views regarding this issue clear:

But to restrict thinking to the patterns merely of English [...] is to lose a power of thought which, once lost, can never be regained. [...] I believe that those who envision a future world speaking only one language, whether English, German, Russian, or any other, hold a misguided ideal and would do the evolution of the human mind the greatest disservice. (1956, p. 244)

By the same token, he was very conscious of the limitations of linguistic metalanguage which, in his opinion, was useless in any other than a strictly grammatical sense (Whorf & Trager, 1996).

Whorf's contribution to linguistics exceeds by far the scope of this paper, which has presented his position on language insofar as it relates to linguistic relativity. Among Whorf's most notable endeavors in other areas were attempts to decipher Maya and Aztec writing, and an analysis of the Hopi language, as well as numerous theoretical articles on phonotactics, the cryptotype, and language typology. Although some of the notions they advanced were either modified or indeed disproved by later research, it should be borne in mind that Whorf was deprived of the opportunity to develop and perfect his work because he died of cancer at the age of 44. His death and that of his mentors, Franz Boas and Edward Sapir, who died within a few years of Whorf, dealt a blow to the linguistic relativity proposal, which was presented to the world in an unfinished form.

Practical Implications. The current reawakening of interest in linguistic relativity and the expansion of the concept to include bilingualism are living proof that the Sapir-Whorf Hypothesis continues to inspire researchers across the academic spectrum. This is most evident in the outpouring of studies and monographs on the subject that has been observed over the past two decades or so. They are a tribute to Whorf in particular for having opened a debate that required expertise and vision, and for having had the courage of his convictions.

It must be kept in mind, however, that Sapir and Whorf were pioneer researchers with a focus on cultural rather than cognitive phenomena. Consequently, the paradigm established for their work could not meet the requirements of more

controlled studies designed to discover relationships between narrowly defined variables. As has been demonstrated in this paper, Whorf's understanding of thought was often intuitive, while his analysis of data was descriptive and impressionistic, even philosophical at times. This alone diminishes the usefulness of the original and/or any other broadly conceived version of linguistic relativity for current research which is based on detailed cognitive models. As Pederson rightly observes, nowadays:

the question 'Does language influence thought?' is being replaced by a battery of questions about whether a given feature of a specific language influences particular cognitive operations, what the exact cognitive mechanisms are which give rise to this influence, and how we can most precisely characterize the nature of this influence. (2007, p. 1036)

In the light of the above, it becomes apparent that the disagreements over what constitutes *Whorfian effects on cognition* that are common in psycholinguistic circles cannot be conclusively resolved because Whorf did not provide enough information on the subject. In fact, he was only a precursor of what later came to be known as a research trend bearing his name.

Personally, I applaud Cook's (2011, p. 12) comment that there is an infuriating tendency for academics "to debate not the actual issues involved, but their interpretation of the writings of Whorf and Sapir [...], rather like the exegesis of a sacred text." Paradoxically, this paper is no exception. Still, as a researcher, I would much rather concentrate on the real issues. One of them is discussed below.

Linguistic and (non?)-linguistic categorization in tasks using linguistic stimuli: A correlational study

Objectives. According to Lakoff (1987, p. 330), the classic Whorfian argument "that the structure of a language could influence nonlinguistic behavior" should be couched in terms of the question of whether naming is part of (non-linguistic) cognition. The present study investigates this issue by attempting to find out to what extent semantic naming distinctions correlate with categorization patterns in tasks that:

- apply non-linguistic criteria of categorization, e.g. similarity judgments in a free sorting task;
- implement categorization according to lexico-semantic criteria which diverge from those laid down by the words under study.

The existence and strength of significant correlations will show whether the observed linguistic patterns may be construed to reliably reflect the underlying conceptual models. This will be the case when items given the same name in the linguistic task are placed in the same category in the non-linguistic condition (Latkowska, 2013).

Participants. The participants of the study were 30 first-year students of the University of Silesia attending a teacher training course at the English Department. Their level of proficiency in English was assessed with the Oxford Quick Placement Test (2001) and ranged between the B2 ($N=22$) and the C1 ($N=8$) levels. None of them had stayed in an English-speaking environment for more than a month. The students also had elementary to lower intermediate knowledge of German, which they had studied in secondary school and at university. Because the study did not examine cross-language comparisons, it was not considered essential to restrict it to monolinguals or limited bilinguals.

Materials. The materials used in the study included nine scenarios (see below), each of which created a context to activate the targeted meaning and related concept. The scenarios were built around the semantic components of Wierzbicka's (1997) Natural Semantic Metalanguage explications for three Polish friendship terms, i.e. *przyjaciel*, *kolega*, and *znajomy*. Eight scenarios were used in a previous study where they elicited the targeted words consistently (Latkowska, 2013).

Przyjaciel 'best friend'

- S1) We went to school together and lived in the same street. On Saturdays we would first meet in the playground, and then, a few years later, on the tennis court. Now we often go to our local for a chat. There isn't a thing we wouldn't know about each other.
- S4) We often talk on the phone or on the net. Our conversations are very honest and deep; sometimes they remind me of going to confession.
- S9) She/he is one of the few people I trust and often discuss my problems with. I admire his/her experience and disinterested wisdom.

Kolega 'friend'

- S2) For five years, we have been meeting at university where we do the same degree course. We sometimes study for exams together and in our free time, i.e. quite rarely, we go to the cinema.
- S6) For several years we've been going to ski camps together. In fact, all of us started from scratch and had many adventures on the ski slopes and routes. We enjoy skiing together.

S10) We are classmates and often stay after school to do our homework together. There are five of us in all and we enjoy studying together.

Znajomy 'acquaintance'

S3) We meet when walking our dogs and often have a chat while our pets chase each other on the grass. This is how I hear the news about the people living in the area.

S5) We met at a conference where we were seated next to each other at the conference dinner. After an interesting conversation we exchanged business cards.

S8) Our kids are classmates and we often meet at parents' meetings or when collecting them from school. Sometimes, when I have to work overtime, s/he walks my son home for me.

Two distractor items were also included, both with a focus on *neighbor*, i.e. S11 and S7, which read as follows:

S7) Although we live on the same floor we meet once a month or less.

S11) We live in a house on the city outskirts. The couple next door are very friendly and we often help each other. For example, last summer I watered their garden when they went on holiday.

The participants were instructed to write down the word(s) they would use to name their relationship with the individual(s) described in each scenario (Test 1). The relationship was not romantic and they could use the same word several times to refer to different scenarios, or use two different words to refer to the same situation. The scenarios were presented in a randomized order which is indicated by the letter S for scenario and the entry number. The participants were expected to provide the target word in response to the prompt: *This person is a(n) or These people are (my)...*. The questionnaire was implemented in Polish.

In addition, the participants were requested to complete a scenario evaluation test (Test 2, see Appendix 1), which was a verbal categorization task aimed at finding out how the participants perceived the scenarios. This was vital to clarifying the nature of the relation signified by each of the friendship terms. The participants were required to assign a corresponding scenario number to one of the following categories:

- a very close and personal relationship;
- a close but not intense, also professional relationship;
- a purely social relationship;
- a purely professional relationship.

To obtain information about the participants' language history, *The Oxford Quick Placement Test* (2001), paper and pen version, was used, as well as a back-

ground questionnaire. The questionnaire focused on factors such as education level and experience with the L2.

Procedure. The tests were administered in the following sequence:

- the free sorting task;
- the scenario naming test (Test1);
- the scenario evaluation test (Test 2).

In the free sorting task, which tends to be used to assess non-verbal categorization (Malt & Ameel, 2011), the participants were presented with a list of unnumbered scenarios and asked to put them into categories on the basis of their similarity. The instructions for the task read as follows: *Mark with the same letter, i.e. A, B, C, and so on, situations which in your opinion are similar to each other/one another.* The participants were free to form as many categories as they saw appropriate. The order of the scenarios was randomized and the participants were not aware of the purpose of the study. The free sorting task preceded the other two tests which required explicit verbalization. It was hoped that implementing the sorting task first would help avoid drawing the respondents' attention to specific linguistic criteria, thus reducing the extent of subvocal verbalization. A decision was also made not to include triads matching in the study since the binary choice enforced by the structure of the task did not always accurately reflect the categorization choices of the respondents who, if given the chance, might have opted for a different answer (Latkowska, 2009). What is more, the piloting stage for the triads task produced inconsistent results.

The naming test and the scenario evaluation test were implemented as described in the section on materials. The respondents were given 45 minutes to complete all three tests. They were also allowed to sign the test sheets with a fictitious name to ensure anonymity. The background questionnaire and the placement test were completed the following week.

Analysis. To compare the similarity of the participants' categorization choices in the three tasks, the situations that were placed in the same category were combined into pairs. For example, if scenarios 1 and 4 elicited the name *przyjaciół*, they were placed in the verbal category of *przyjaciół* under the label 1, 4. If these two scenarios were marked with the same letter in the free sorting task, they were obviously in the same category, too, which was considered to be a measure of similarity between verbal categorization and sorting behavior. For the sake of precision, the pairing of the situations from the sorting test was carried out using an Excel macro created for this purpose. The pairs were then tallied and the score for each pair was correlated with those for the equivalent pairs in the other two tasks. The mode of analysis developed for this study is partly modeled on Malt et al. (1999). The Shapiro-Wilk test was run to assess the distribution of the data. Since they did not follow normal distribution

($W > 0.19$, $p = 0.000$) in all three tests, non-parametric Spearman rank-order correlations were computed. The alpha level was set at 0.05 or less.

Results and Discussion. In the free sorting task, the subjects created six categories marked from A to F. Category F contained just one scenario (S7) and was subsequently excluded from the analysis. The sort yielded 42 pairings of scenarios, the most frequent of them being 4, 9 ($N = 30$), 1, 4 ($N = 25$), 1, 9 ($N = 25$), 8, 11 ($N = 23$), 10, 2 ($N = 20$), 2, 6 ($N = 16$), and 5, 7 ($N = 16$).

The naming task elicited four main groupings, i.e. (1) *przyjaciel* ($N = 84$), (2) *kolega* ($N = 83$), (3) *znajomy* ($N = 108$) and (4) *sąsiad* ($N = 45$). There were also infrequent cases of other names. All of the scenarios elicited their targets, which in most cases constituted about 70% of each category's elicited name set, and were thus the dominant names. As regards the scenarios that received the same name, there were 38 pairings altogether. The most frequent ones included: 4, 9 ($N = 26$), 1, 4 ($N = 24$), 1, 9 ($N = 21$), 10, 2 ($N = 21$), 10, 6 ($N = 21$), 5, 8 ($N = 21$), 3, 5 ($N = 19$), and 7, 11 ($N = 19$).

Test 2 served as a control for the naming task and produced a clear response pattern for S1, 4, and 9 (*przyjaciel* 'close friend'), which were classified as *very close and personal* by between 25 and 30 respondents (83%–100%). S5 for *znajomy* was evaluated as *purely professional* ($N = 25$) by the majority of respondents, too. There was more variability in the responses to the remaining situations. Overall, the test produced 38 scenario pairings, the most frequent being: 4, 9 ($N = 27$), 1, 9 ($N = 25$), 1, 4 ($N = 23$), 10, 2 ($N = 20$), 10, 8 ($N = 12$), 2, 6 ($N = 12$), 2, 8 ($N = 11$) and 2, 11 ($N = 11$).

Spearman rank-order correlations computed for specific categories in the free sorting and naming tests showed clearly that the two tests summoned different categorization criteria. The sorting task (Table 2, see Appendix 3) elicited judgments based on the most salient and extreme properties, such as very intense and personal contact (S1, 4, 9) on the one hand, and little or no contact (S5, 7) on the other. Situations involving relationships of moderate intensity were perceived as similar.

In the naming task (Table 1, see Appendix 2), *znajomy* covered situations involving varying degrees of intensity, except for the most powerful ones, which were assigned primarily to *przyjaciel*, and secondarily to *kolega*. There were no statistically significant positive correlations between the words. Thus, although similar at first glance, the patterns of similarity identified in the naming task did not match. Neither did the criteria applied in the sorting test reflect those exhibited in the naming task.

A between-test comparison involving the sorting test and the scenario evaluation test (Test 2) yielded surprising between-task synchrony (Table 2, see Appendix 3), with Grouping A correlating significantly with categories evaluated as *very close and personal* ($\rho = 0.62$, $p = 0.000$), Grouping B showing

similarity to *close but not intense* relationships ($\rho = 0.41$, $p = 0.004$), Grouping C aligning with *purely social* relations ($\rho = 0.36$, $p = 0.01$), and Grouping D correlating significantly with *purely professional* scenarios ($\rho = 0.36$, $p = 0.01$). Grouping E correlated with *purely social* relations as well ($\rho = 0.30$, $p = 0.04$).

The categories of the naming test and those of the evaluation test displayed a comparable level of similarity (Table 1, see Appendix 2). Namely, *przyjaciel* was strongly correlated with the *very close and personal* parameter ($\rho = 0.72$, $p = 0.000$), *kolega* correlated with the *close but not so intense* category ($\rho = 0.58$, $p = 0.000$), while *znajomy* was aligned to both *close but not so intense* and *purely social* relations ($\rho = 0.3$ and $\rho = 0.44$, respectively, $p < 0.05$). The significant correlate for *sąsiad* ‘neighbor’ was *purely social* relations ($\rho = 0.50$, $p = 0.000$). These results are too systematic to be accidental and indicate that semantic categories are the most precise medium of description for other semantic categories.

Conclusions. The results of this small-scale study show that language-based tests may yield quite misleading results when used to examine non-verbal categorization patterns. Although the data demonstrate an obvious dissociation of lexical naming in Test 1 from the allegedly non-linguistic criteria applied in the free sorting task, the dissociation is called into question by the correlations obtained for the sorting task and the scenario evaluation test (Test 2). Because these are so consistent they raise the possibility that more general linguistic criteria were involved in the sorting process. This does not seem unlikely since linguistic stimuli activate entire language systems, bringing to bear context-relevant parameters and criteria. It is their interfering influence that might have cancelled out the effects of the naming distinctions from Test 1. As things stand, however, we have no way of knowing whether the observed trends reflect deeper conceptual distinctions or the dynamics of semantic processing at the linguistic level.

This lack of clarity is of relevance to research into all things Whorfian, where it is necessary to distinguish between the conceptual and linguistic levels of representation. Since language-based tests allow for the involvement of untargeted linguistic categories, which may affect the validity of findings, their usefulness for research into non-linguistic categorization should be seriously questioned.

As a concluding remark, it might be worth noting that thanks to the application of sophisticated research technologies such as, e.g. fMRI and ERPs, researchers are now studying the language-cognition interface in ways that were unimaginable to Whorf and his contemporaries. There can be no doubt that, despite many shortcomings, the Sapir-Whorf Hypothesis continues to fascinate the academic community, inspiring further research into the unsolved mysteries of the human mind.

Polish version of Test 2

Które z przedstawionych sytuacji odnoszą się do relacji (można pominąć niektóre z podanych poniżej punktów)

Bardzo bliskich i osobistych (podaj numery sytuacji)

.....
.....
.....
.....
.....

Zażytych, lecz mniej osobistych, również zawodowych (podaj numery sytuacji)

.....
.....
.....
.....
.....

Typowo towarzyskich (podaj numery sytuacji)

.....
.....
.....
.....
.....

Typowo zawodowych (podaj numery sytuacji)

.....
.....
.....
.....
.....

Appendix 2

Table 1
Spearman's correlation coefficients for the linguistic tests

N = 48	Przyjaciel	Kolega	Znajomy	Sąsiad	Very close	Close but not intense	Social	Professional
Przyjaciel rho	n/a	$\rho = -0.17$ $p = 0.25$	$\rho = -0.39$ $p = 0.005$	$\rho = -0.09$ $p = 0.51$	$\rho = 0.72$ $p = 0.000$	$\rho = -0.46$ $p = 0.001$	$\rho = -0.19$ $p = 0.19$	$\rho = -0.07$ $p = 0.59$
Kolega rho	$\rho = -0.17$ $p = 0.25$	n/a	$\rho = 0.17$ $p = 0.24$	$\rho = -0.21$ $p = 0.13$	$\rho = -0.12$ $p = 0.41$	$\rho = 0.58$ $p = 0.000$	$\rho = -0.13$ $p = 0.35$	$\rho = -0.03$ $p = 0.83$
Znajomy rho	$\rho = -0.39$ $p = 0.005$	$\rho = 0.17$ $p = 0.24$	n/a	$\rho = 0.25$ $p = 0.08$	$\rho = -0.46$ $p = 0.000$	$\rho = 0.30$ $p = 0.04$	$\rho = 0.44$ $p = 0.001$	$\rho = 0.25$ $p = 0.08$
Sąsiad rho	$\rho = -0.09$ $p = 0.51$	$\rho = -0.21$ $p = 0.13$	$\rho = 0.25$ $p = 0.08$	n/a	$\rho = -0.11$ $p = 0.43$	$\rho = 0.006$ $p = 0.96$	$\rho = 0.50$ $p = 0.000$	$\rho = -0.05$ $p = 0.71$

Appendix 2

Table 2
Spearman's correlation coefficients for the free sorting task and the linguistic tests

N = 48	Grouping A	Grouping B	Grouping C	Grouping D	Grouping E	Przyjaciel	Kolega	Znajomy	Sąsiad	Very close and personal	Close but not intense	Social	Professional
Grouping A	$\rho = 0.086$	$\rho = 0.086$	$\rho = 0.072$	$\rho = -0.003$	$\rho = -0.356$	$\rho = 0.602$	$\rho = 0.292$	$\rho = -0.254$	$\rho = -0.243$	$\rho = 0.617$	$\rho = 0.058$	$\rho = -0.188$	$\rho = -0.033$
Fig. (2-tailed)	$p = 0.562$	$p = 0.627$	$p = 0.627$	$p = 0.983$	$p = 0.013$	$p = 0.000$	$p = 0.044$	$p = 0.081$	$p = 0.096$	$p = 0.000$	$p = 0.696$	$p = 0.201$	$p = 0.786$
Grouping B	$\rho = 0.086$	n/a	$\rho = 0.636$	$\rho = 0.447$	$\rho = 0.056$	$\rho = -0.006$	$\rho = 0.273$	$\rho = 0.419$	$\rho = 0.154$	$\rho = -0.091$	$\rho = 0.409$	$\rho = 0.245$	$\rho = -0.181$
Fig. (2-tailed)	$p = 0.562$	$p = 0.000$	$p = 0.000$	$p = 0.001$	$p = 0.703$	$p = 0.966$	$p = 0.060$	$p = 0.003$	$p = 0.295$	$p = 0.538$	$p = 0.004$	$p = 0.093$	$p = 0.218$
Grouping C	$\rho = 0.072$	$\rho = 0.636$	n/a	$\rho = 0.487$	$\rho = -0.032$	$\rho = 0.093$	$\rho = 0.024$	$\rho = 0.331$	$\rho = 0.252$	$\rho = -0.024$	$\rho = 0.209$	$\rho = 0.360$	$\rho = 0.024$
Fig. (2-tailed)	$p = 0.627$	$p = 0.000$	$p = 0.000$	$p = 0.000$	$p = 0.831$	$p = 0.528$	$p = 0.871$	$p = 0.021$	$p = 0.083$	$p = 0.870$	$p = 0.155$	$p = 0.012$	$p = 0.873$
Grouping D	$\rho = -0.003$	$\rho = 0.447$	$\rho = 0.487$	n/a	$\rho = 0.198$	$\rho = 0.158$	$\rho = -0.114$	$\rho = 0.387$	$\rho = -0.011$	$\rho = 0.081$	$\rho = -0.121$	$\rho = 0.225$	$\rho = 0.359$
Fig. (2-tailed)	$p = 0.983$	$p = 0.001$	$p = 0.000$	$p = 0.176$	$p = 0.176$	$p = 0.282$	$p = 0.441$	$p = 0.007$	$p = 0.939$	$p = 0.585$	$p = 0.413$	$p = 0.124$	$p = 0.012$
Grouping E	$\rho = -0.356$	$\rho = 0.056$	$\rho = -0.032$	$\rho = 0.198$	n/a	$\rho = -0.142$	$\rho = -0.196$	$\rho = 0.292$	$\rho = 0.422$	$\rho = -0.168$	$\rho = -0.215$	$\rho = 0.295$	$\rho = 0.262$
Fig. (2-tailed)	$p = 0.013$	$p = 0.703$	$p = 0.831$	$p = 0.176$	$p = 0.176$	$p = 0.335$	$p = 0.182$	$p = 0.044$	$p = 0.003$	$p = 0.253$	$p = 0.143$	$p = 0.042$	$p = 0.057$

References

- Athanasopoulos, P. (2011a). Linguistic relativity in SLA. Thinking for speaking (book review). *International Journal of Bilingual Education and Bilingualism*, 14(5), 621–625.
- Athanasopoulos, P. (2011b). Cognitive restructuring in bilingualism. In A. Pavlenko (Ed.), *Thinking and speaking in two languages* (pp. 29–65). Bristol: Multilingual Matters.
- Bylund, E. (2011). Language-specific patterns in event conceptualization: Insights from bilingualism. In A. Pavlenko (Ed.), *Thinking and speaking in two languages* (pp. 108–142). Bristol: Multilingual Matters.
- Carroll, J. B. (1956). Introduction. In J. B. Carroll (Ed.), *Language, thought, and reality. Selected writings of Benjamin Lee Whorf* (pp. 1–34). Cambridge, MA: The M.I.T. Press.
- Casasanto, D. (2008). Who's afraid of the big bad Whorf? Crosslinguistic differences in temporal language and thought. *Language Learning*, 58: Suppl. 1, 63–79.
- Cook, V. (2011). Relating language and cognition: The speaker of one language. In V. Cook and B. Bassetti (Eds.), *Language and bilingual cognition* (pp. 3–22). New York and Hove: Psychology Press.
- Cook, V., & Bassetti, B. (2011). Relating language and cognition: The second language user. In V. Cook & B. Bassetti (Eds.), *Language and bilingual cognition* (pp. 143–190). New York and Hove: Psychology Press.
- Gendron, M., Barsalou, L., Lindquist, K. A., & Feldman B. L. (2012). Emotion words shape emotion percepts. *Emotion*, 12(2), 314–325.
- Goddard, C. (2003). Whorf meets Wierzbicka: Variation and universals in language and thinking. *Language Sciences* 25, 393–432.
- Han, Z., & Cadierno, T. (2010). Preface. In Z. Han & T. Cadierno (Eds.), *Linguistic relativity in SLA. Thinking for speaking* (pp. xi–xv). Clevedon, UK: Multilingual Matters.
- Kramsch, C. (2004). Language, thought and culture. In A. Davis & C. Elder (Eds.), *Handbook of Applied Linguistics* (pp. 235–261). Oxford: Blackwell Publishing.
- Lakoff, G. (1987). *Women, fire and dangerous things: What categories reveal about the mind*. Chicago: Chicago University Press.
- Latkowska, J. (2009). Conceptual transfer from a bilingual perspective: Research trends and suggestions. *Linguistica Silesiana*, 30, 239–252.
- Latkowska, J. (2013). *The language-cognition interface in bilinguals: An evaluation of the Conceptual Transfer Hypothesis*. Katowice: University of Silesia Press.
- Levinson, S. (2003). *Space in language and cognition*. Cambridge: Cambridge University Press.
- Lucy, J. (2004). Language, culture and mind in comparative perspective. In M. Achard & S. Kemmer (Eds.), *Language, culture and mind* (pp. 1–22). Stanford, CA: CSLI Publications.
- Malt, B. C., Sloman, S., Gennari, S., Shi, M., & Wang, Y. (1999). Knowing versus naming: Similarity and the linguistic categorization of artifacts. *Journal of Memory and Language*, 40, 230–262.
- Malt, B., & Ameel, E. (2011). The art and science of bilingual object naming. In A. Pavlenko (Ed.), *Thinking and speaking in two languages* (pp. 170–197). Clevedon, UK: Multilingual Matters.
- Pavlenko, A. (2014). *The bilingual mind and what it tells us about language and thought*. Cambridge: Cambridge University Press.
- Pavlenko, A., & Malt, B. (2011). Kitchen Russian: Cross-linguistic differences and first-language object naming by Russian-English bilinguals. *Bilingualism: Language and Cognition*, 14(1), 19–45.

- Pederson, E. (2007). Cognitive linguistics and linguistic relativity. In D. Geeraerts & Cuyckens, H. (Eds.), *The Oxford Handbook of Cognitive Linguistics* (pp. 1012–1245). Oxford: Oxford University Press.
- Slobin, D. I. (1996). From “thought and language” to “thinking for speaking.” In J. J. Gumperz & S. C. Levinson (Eds.), *Rethinking linguistic relativity* (pp. 70–96). Cambridge: Cambridge University Press.
- Slobin, D. I. (2003). Language and thought on line: Cognitive consequences of linguistic relativity. In D. Gentner & S. Goldin-Meadow (Eds.), *Language and mind. Advances in the study of language and thought* (pp. 157–192). Cambridge, MA: The MIT Press.
- Slobin, D. I. (2004). The many ways to search for a frog: Linguistic typology and the expression of motion events. In S. Strömquist & L. Verhoeven (Eds.), *Relating events in narrative: vol. 2. Typological and contextual perspectives* (pp. 219–257). Mahwah, NJ: Lawrence Erlbaum.
- Slobin, D. I. (2005). Linguistic representations of motion events: What is signifier and what is signified? In C. Maeder, O. Fischer, & W. J. Herlofsky (Eds.), *Iconicity inside out: Iconicity in language and literature, 4* (pp. 307–322). Amsterdam/Philadelphia: John Benjamins.
- Stutterheim, Ch. von (2003). Linguistic structure and information organization: The case of very advanced learners. *EUROSLA Yearbook, 3*, 183–206.
- Whorf, B. L. (1956). On psychology. In J. B. Carroll (Ed.) *Language, thought, and reality. Selected writings of Benjamin Lee Whorf* (pp. 40–42). Cambridge, MA: The MIT Press.
- Whorf, B. L. (1956). A linguistic consideration of thinking in primitive communities. In J. B. Carroll (Ed.), *Language, thought, and reality. Selected writings of Benjamin Lee Whorf* (pp. 65–86). Cambridge, MA: The MIT Press.
- Whorf, B. L. (1956). The relation of habitual thought and behavior to language. In J. B. Carroll (Ed.), *Language, thought, and reality. Selected writings of Benjamin Lee Whorf* (pp. 134–159). Cambridge, MA: The MIT Press.
- Whorf, B. L. (1956). Science and linguistics. In J. B. Carroll (Ed.), *Language, thought, and reality. Selected writings of Benjamin Lee Whorf* (pp. 207–219). Cambridge, MA: The MIT Press.
- Whorf, B. L. (1956). Linguistics as an exact science. In J. B. Carroll (Ed.), *Language, thought, and reality. Selected writings of Benjamin Lee Whorf* (pp. 220–231). Cambridge, MA: The MIT Press.
- Whorf, B. L. (1956). Languages and logic. In J. B. Carroll (Ed.), *Language, thought, and reality. Selected writings of Benjamin Lee Whorf* (pp. 233–245). Cambridge, MA: The MIT Press.
- Whorf, B. L. (1956). Language, mind, and reality. In J. B. Carroll (Ed.), *Language, thought, and reality. Selected writings of Benjamin Lee Whorf* (pp. 246–270). Cambridge, MA: The MIT Press.
- Whorf, B. L., & G. L. Trager (1996). The Yale Report. In P. Lee (Ed.), *The Whorf Theory Complex* (pp. 251–280). Amsterdam/Philadelphia: John Benjamins.
- Wierzbicka, A. (1997). *Understanding cultures through their key words*. Oxford: Oxford University Press.

Jolanta Latkowska

Wie wichtig ist die Sapir-Whorf-Hypothese für gegenwärtige psycholinguistische Forschungen?

Zusammenfassung

In dem Artikel möchte die Verfasserin die Frage beantworten, ob die Sapir-Whorf-Hypothese ein Bezugspunkt für gegenwärtige Forschungen über den Zusammenhang zwischen der Sprache und dem kognitiven Apparat des Menschen (dem Denken) werden kann. Die Diskussion fußt auf Benjamin Lee Whorfs (1956) Werken; berücksichtigt werden dabei die für die oben genannte Hypothese wichtigsten Begriffe: Sprache, Gedanke und Verhalten. Gesucht wird auch die Antwort auf andere Frage: ob Sprachkategorien ein gewisses Fenster zu den zu Grunde liegenden Denkprozessen sind, wie Whorf und einige von den heutigen Tendenzen der kognitiven Sprachwissenschaft behaupten? Bei der Analyse bedient sich die Verfasserin der in verbalen und nonverbalen Texten angewandten Kategorisierungskriterien.