Original Papers

Polish Psychological Bulletin
2009, vol. 40 (4), 15-20
DOI – 10.2478/s10059-009-0012-4

*Monika Malinowska**
*Maciej Haman**

Subfocal Color Categorization and Naming: The Role of Exposure to Language and Professional Experience

The current state of the debate on the linguistic factors in color perception and categorization is reviewed. Developmental and learning studies were hitherto almost ignored in this debate. A simple experiment is reported in which 20 Academy of Fine Arts, Faculty of Painting students' performance in color discrimination and naming tasks was compared to the performance of 20 Technical University students. Subfocal colors (different hues of red and blue) were used. While there was no difference in overall discrimination ability, AFA students had a much richer and specialized color vocabulary. Both groups also applied different strategies of discrimination and naming. However, naming system in neither group was coherent. This suggests that naming played primarily the role of markers for control processes rather than names for categories. It is concluded that up-to-date debate is too simplified and a complex model of interrelations between perceptual categorization and naming framed in the developmental context is needed rather than the search for a simple answer "language", "environment", or "perceptual universals".

Key words: *color perception, color categorization, linguistic relativity hypothesis, universalist hypothesis, cognitive development*

Color perception, categorization and naming are in the focus of the debate between culturalist-relativists' (Whorfians') and universalists' positions in the studies of human cognition. The argument in this field began with Brown and Lenneberg's (1954) work. Their experiment showed that colors which are named in a given language by the same term are judged to be more similar than colors representing the same physical distance but named with a different term. Their results had set the counter on the relativist side until Berlin and Kay's (1969) systematic study of color terms in over twenty languages (extended later onto almost one hundred languages in the World Color Survey). They found that, although languages differ in number and scope of color words, these differences are highly systematic. If there are only two color terms they refer to white and black, in the next step red is added, then green and yellow, and so on till the maximum of eleven basic (focal) color terms is reached, which can be found in many languages, including Polish and English. Berlin and Kay (1969) assume that these eleven focal colors provide a universal perceptual system of categories that constrains the variety of culture- and language-dependent color conceptualization. Cross-linguistic and cross-cultural studies by Eleanor Heider-Rosch and her colleagues (Heider & Olivier, 1972; Rosch, 1972) demonstrated that even in culture that has a very spars system of color terms, learning categories and memory for color distinctions that respect focal color boundaries is significantly better than for colors with arbitrarily set boundaries (with hue, brightness and saturation differences kept constant). In the over fifty years since Brown and Lenneberg's study, research in cognition and cognitive development has revealed such enormous progress that the original problem of linguistic relativity seems naïve and oversimplified, and color perception seems to be one of the most misleading areas in this argument, yet it bursts forth with new force from time to time, mostly driven by history and ambition. Its current stage was initiated by Davidoff, Roberson, and Davis (1999), and Roberson,

* **Faculty of Psychology, University of Warsaw; e-mail: malinowskaaa@gmail.com, meh@psych.uw.edu.pl**

Davis, and Davidoff's (2000) publications on Papua New Guinea tribal culture, Berinmo, whose language uses only five color terms.

First of all, Berinmo language does not respect focal color boundaries revealed in World Color Survey. This could suggest that the only perceptual constraint over color classification is similarity: color categories should cover a continuous and relatively compact area in the color space defined on the dimensions of hue and brightness. Moreover, Roberson, Davis, and Davidoff (2000) attempted to replicate Heider and Olivier's (1972) results in Berinmo, using highly saturated focal colors that are represented in English but not in the Berinmo language ("good examples"), and their low saturation hues ("poorer examples"). They found only very weak effects of "goodness" of color on the memory performance, and, although the results were in general in conformity with those of Heider and Olivier, a careful control over the experimental procedure revealed that even these weak effects are likely to be products of experimental artifacts.

But again a closer look into these results makes the argument much less clear. Kay and Regier (2006) showed through a formal analysis that the Berinmo system of color names gives the best possible fit to focal colors if the category size is held constant. And Munnich and Landau (2003) argued that Roberson et al.'s (2000) procedure was highly language-dependent itself. Subjects were engaged in overt rehearsal of color names, and that could erase any pre-linguistic categorical effects.

So far, the developmental issues were almost ignored in this debate. It was well known that color names are learned slowly, with errors and large individual variability in the early stages of this process (see e.g. Soja, 1994). That was interpreted in line with the Whorfian position (there is no pre-linguistic color categorization system that can facilitate color-term learning), as well as against it (color words are not indispensable to differentiate and use color categories, and the linguistic entry to color cognition is not a natural one). It was also known, although often contested, that categorical perception of colors can be found even in infants (Borstein, 1985). But for example studies of the use of color information in object individuation and categorization, which precedes learning of color terms by a few years (e.g. Wilcox, 1999), were absent in the debate. Also, other relevant knowledge from the still widening area of developmental neurocognitive science was rarely called into the debate.

This situation is beginning to change. Impressive studies by Gilbert et al. (2006) have revealed the complexity of the relation between language and color categorization at the neural level. Gilbert and his colleagues administered to their subjects a color discrimination task. Targets were either in the same nominal color category or in a different one than comparison examples, and were presented only either to the left or to the right visual field. A categorical perception effect (quicker cross-category differentiation) was found only in the right visual field, and was erased with overload of verbal (but not visual) working memory. This means that linguistic categories affect only the left hemisphere color representation sub-system. However, some results of the study could suggest an even further going interpretation. Overload of verbal memory also increased reaction times for targets presented to the left visual field, and although reaction times did not reveal clear categorical perception effect (mean difference in RTs for within-category and cross-category targets was relatively large, but not significant), they were exactly the reverse of those for the targets presented to the right visual field. Perhaps when the language-controlled categorical perception in the left hemisphere was blocked by working memory overload, nonlinguistic right hemisphere system took over the categorization task. It seems even more likely in the light of Franklin et. al.'s (2008) developmental study, attempting to replicate Gilbert et al.'s results in toddlers at different stages of the acquisition of color terms. While in subjects in early stages of learning color terms the categorical perception effects were found mostly in the left visual field, in competent color term users the effect was found in the right visual field (like in adults in the Gilbert et al.'s study). Several other important recent contributions to the field of the development of color perception and naming, including Roberson et al. (2004) and Franklin and Davis (2004), will not be discussed here in detail (for more detailed reviews and discussion see e.g. Carruthers, in press; Regier and Kay, 2009; Roberson, 2005). All this evidence together shows a complex, dynamic interrelation between perceptual constraints and the color naming system, that is based on both competition and collaboration between them.

The study presented in this paper is designed to investigate the role of language in learning color classification after acquisition of basic color terms in adults who need fine color categories in their professional activities. Studying such a group seems to provide an interesting contribution to the field through investigating the differences within the same linguistic group, testing if the learning process changes when the basic system of color categories is acquired, and introducing an element of color-related activity, which was also supposed to be important but was rarely studied. Granting the truth of the contemporary version of the relativist position, we can assume that once language has reached control over the color categorization system it should be an effective tool for expanding this system. Everywhere a finer-than-common classification of colors is needed for pragmatic and communicative purposes, and where the social environment offers a specialized, professional lexicon, this tool should be easily used to tune perceptual discrimination as well as to increase the categorical distinctiveness, at least as far as the most general perceptual constraints (focal category borders, general perceptual thresholds, similarity) are satisfied. To test this hypothesis we have identified a professional group that needs fine color classification and is supposed to have a rich and specialized system of color names (i.e. painters), and we recruited from this subpopulation a sample that should represent a pre-expert, middle-experienced level (i.e. students of the Academy of Fine Arts, Faculty of Painting). This group was compared with another one that does not need sophisticated categorization of colors in their activities (Technical University Students) on the performance of color discrimination and categorization tasks.

Method

Participants

Forty university-level students participated. Twenty of them were recruited at the Faculty of Painting, Warsaw Academy of Fine Arts (AFA) and the remaining twenty at the Faculty of Electronics and Information Technology, Warsaw University of Technology (TU) (gender proportion was roughly equated in both groups).

Materials

Materials consisted of three A4-size sheet booklets. The first sheet contained the instruction, the second one a discrimination task, and the third one a naming task (the order of discrimination and naming tasks was rotated).

Discrimination task. Ten randomly selected fields of the six by six squares (2 cm x 2 cm each) chart were filled with colored circles. There were two variants of the color chart: a red and a blue one. In both cases the colors of the circles were constructed by selecting the most saturated hue of either red or blue and softening it in nine steps till 30 per cent saturation. The participants' task was to order the circles from the darkest to the brightest one. They were not informed that every shade is different, and so they were allowed to use tied ranks.

Naming task. Five colored circles were printed row by row (in the random order) on the left side of the page. The colors were selected from those used in the discrimination task, by taking the initial (the most saturated) one, and every even-step transformation. The participants' task was to name each of the colors (number of words was not restricted). Those subjects who got the blue version of discrimination task got the red version of naming task and vice versa.

Procedure

Subjects were tested in bright natural light conditions. They were asked to participate in a study of color perception and given the booklets. In an instruction on the first page they were asked to write down in which faculty they are studying, and then to read carefully the instructions, and to perform two separate tasks on the two following pages. Time was not limited, however the entire task never took more than five minutes.

Results

Discrimination Task

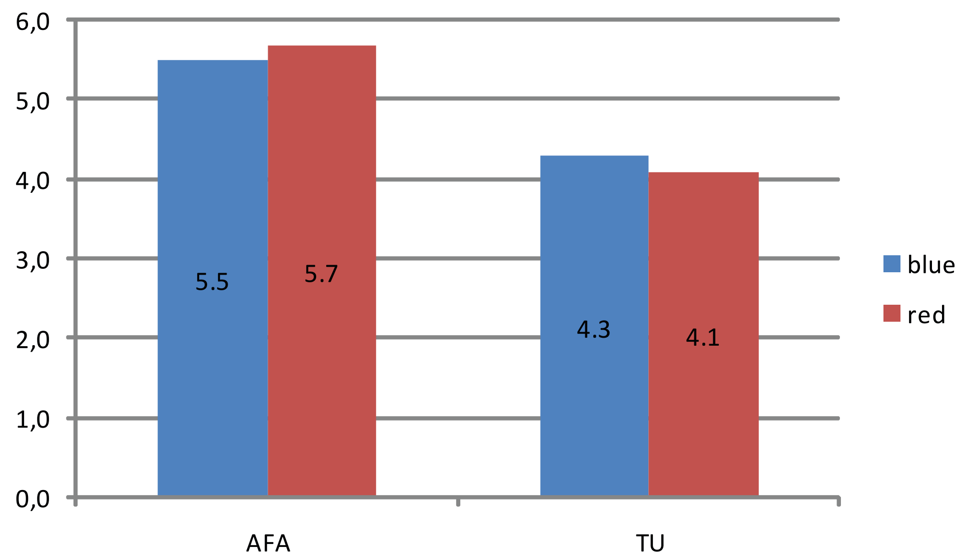
Scoring system and evaluation of discrimination task. Participants' rankings were transformed onto a ten point scale (tied ranks were calculated as the mean of the appropriate numbers of subsequent ranks), and the mean of differences between ideal and subjects' rankings were computed as a measure of the accuracy (error level) of color discrimination. The task seems to be well calibrated. It was not too hard, and every participant performed above random level. However, it was hard enough to differentiate participants (the scores ranged from 0 to .7, while the random level was 2.5). Only four out of 40 participants made no error.

Between-group comparison. 2 (faculty) x 2 (color) between-group ANOVA was run on discrimination scores. Neither faculty, color, or interaction of both even approached significance (all $F < 1$). Mean scores for the Academy of Fine Arts and the Technical University students were almost indistinguishable: ($M = 0.255$, $SD = 0.182$, and $M = 0.260$, $SD = 0.198$, respectively). That means that the experience with colors in painting activity did not affect real color discrimination accuracy. However, one interesting difference could be noted: TU students more often assigned transposed ranks (seven out of 20 TU students, and only one out of 20 AFA student made at least one transposition error, the difference is significant, *exact Fisher* $p = 0.0436$), while AFA students more often used tied ranks (mean numbers of separate rank levels were $M = 7.55$, $SD = 1.79$ for AFA students, and $M = 8.15$, $SD = 1.98$ for TU students), although this difference was not significant. That seems to reflect task performance strategy rather than perceptual difference. In uncertain cases, TU students made finer, but sometimes erroneous discriminations, while AFA students more often reported lack of differences (tied ranks).

Naming Task

The number of different color names (tokens) produced in response to five test colors was determined for all participants. Again, 2 (faculty) x 2 (color) between-group ANOVA was run. There was a highly significant effect of the faculty ($F(1,36) = 19,82$, $p < 0,001$), and no effect of color or interaction (both $F < 1$). AFA students produced

Figure 1. Mean numbers of different color names produced by each participant (by group and color)



more color names (on average) as TU students (see Figure 1 for illustration of this effect). A difference could be also seen when the total number of different color names is compared between groups. AFA students produced 62 different color names (32 for blue and 30 for red) while TU students produced only 43 (18 for blue, 25 for red).

Qualitative analysis of the color names produced by the two group also suggests some interesting between group differences (see Table 1). The names produced by TU students are in the majority either standard focal color names that were further modified by two kinds of additions: either words or prefixes that refer to saturation and brightness (e.g. "ciemnoniebieski" – dark blue), or referred to real object color designates ("sok z buraków" – red beetroot juice). The same naming strategies could be also found in AFA students, but they also used more specific, technical names (like 'indigo', 'cobalt', 'Paris blue').

But if coherence of color name extensions is considered, it could be shown that these specific names are used inconsistently both within- and between-subjects. Almost any of the names that were used more than three times referred to hues that were separated by one or more intermediate shades (the only exemption was 'Prussian blue', which was used to name two neighboring colors). Moreover, an interesting strategy could be observed in at least 34 subjects (14 AFA students and 10 TU students). The same term (sometimes with a different modifier) was used to name nonadjacent hues while another term was used to name intermediate colors (e.g. 'blue – ultramarine – blue – dark blue' or 'turquoise - cobalt – blue – turquoise'). This strategy could suggest that there is a need for linguistic markers of perceptual differences, but there is no coherent color categorization system below focal color level to which naming could be bootstrapped. This is surprising at least in AFA students, as they should be intensively exposed to fine color categorization, and there should be at least some consensus concerning professional terms for the color palette.

Table 1. Number of tokens in five lexical categories (the same terms used by different subjects were counted separately)

		Simple word - focal	Focal with modifier (common)	Focal with modifier (technical)	Simple - common	Simple - technical
	red	13	22	6	4	4
AFA	blue	7	14	9	1	16
	Total AFA	**20**	**36**	**15**	**5**	**20**
	red	23	20	0	4	1
TU	blue	26	17	0	3	2
	Total TU	**49**	**37**	**0**	**7**	**3**
TOTAL		**69**	**73**	**15**	**12**	**23**

Discussion

Our results show that neither intensive exposure to culturally determined activities requiring fine perceptual color discrimination, nor exposure to a rich system of color names, affects color perception, although they lead to a significant increase in the subject's color vocabulary size. However, at least within the focal color categories, this vocabulary is not coherently mapped onto any system of subcategories. It seems that the process of learning fine color distinctions at the lexical level is like the first steps in color naming: new below-focal level color terms are learned slowly and used inconsistently, despite intensive training and appropriate discriminative abilities at the perceptual level.

So far these data are in clear opposition to the original form of the Whorf hypothesis, and are only barely compatible with its softer contemporary forms. However, this is not a position we endorse here. We would like to argue that the whole problem of linguistic determinism/relativism in color perception is misstated. Let us summarize some recent findings mentioned in the introductory part of this paper. It is now well established that color perception is categorical, and that consistent categorical color perception and use of color categorization in other cognitive tasks can be observed at least at the end of the first year of life, a few years before acquisition of color names. The process of acquisition of the color terms itself is slow and starkly different from e.g. learning common names for objects or verbs referring to everyday activities. On the other hand, the neural basis of the color categorization changes with the acquisition of color terms. Old neural circuits remain active, however, and process the input from the left visual field. The majority of cultures and languages use a categorical system based on universal focal colors with a less constrained system of subcategories within them. There are also cultures that use very sparse and sometimes odd systems of lexical color categories. Some weak effects of the universal perceptual-categorical system could be found also in these cultures. All that suggests that the interrelation between color perception, categorization, and language is a complex and dynamic one[1]. Simple declaration like "there are some perceptual constraints, but cultural-linguistic categorization could be overwritten over them" cannot describe appropriately this interrelation. Perceptual-categorical universals provide more than starting point to word learning, and language can affect categorization in many ways, not only through forcing culture-dependent ready-made category system, as suggested by the typical instantiations of the relativity hypothesis. Firstly, through lexicalization of color categories children gain a powerful representational tool. Named color categories, through magneto effect, become sharper and better crystallized. Secondly, language provides tools for communicating about colors, and through communicative pragmatics can direct attention to focus on some culture-relevant focal and finer-than-focal level perceptual distinctions, or to ignore these that are irrelevant even if are made at the basic level. Thirdly, language plays a special role in developing cognitive control system (executive function), as was proven in several studies, also such that largely employed color classification, e.g. in the card sorting task or Stroop test (see e.g. Diamond, 2006). We can speculate that this was the function of language responsible for the between-group differences found in the present study. AFA students, whose activities and communication needs require better control over color processing, had richer color-related vocabulary, and employed different perceptual and naming strategies. Their perceptual discrimination was more careful (which probably marks higher executive function involvement) – they used tied ranks rather than make transposition errors. They also made more lexical distinctions, but these distinctions were often surprising if regarded from the perspective of a coherent categorization system: participants in our study tended to use different names for neighboring colors, but often named distant colors with the same term. While this is odd from the point of view of the similarity-based theory of categorization, it seems to be reasonable if naming plays primarily the role of the system of markers used by the control processes. But although language severely affects categorization in so many interrelated ways, there is no proof that it is a source of categorization, or that it determines categories in any deep sense. On the contrary, it seems that the process of mapping color names onto color categories is slow, incoherent and leads to many errors at the early stages, and that the same happens in adults when required to learn some finer color distinctions (e.g. AFA students in our study; this is also a possible explanation for only weak effects of teaching new color distinctions in cultures with poor color-related language; e.g. Roberson et al., 2000).

Unfortunately many researchers, especially on the Whorfian side, still restrict the debate to the level of simple solutions, following some historical reminiscences. There are only a few exceptions, including the recent review papers by Roberson (2005) or Regier and Kay (2009).

This situation is striking in comparison to the analogous area of spatial cognition, which is also a field of the universalist – relativist debate. Spatial terms, like colors names, are also acquired relatively late, and used inconsistently at the early stage. On the other hand, even very young children seem to have developed a system of spatial categories. Acquiring spatial lexicon affects both classification of spatial relations and controlling over processing

[1] We have left out of the debate the environmental factors, like color distribution in a given environment, or an exposition to ultraviolet light, which are also claimed to shape color perception and the system of color categories (see e.g., Shepard, 2001). But see also Steels and Belpaeme (2005) for a formal model that shows that environmental factors alone are not sufficient to form the color categorization system without any linguistic and perceptual constraints.

spatial information. So the analogy seems to be deep, but the research strategies radically differ. At least for the last twenty years or so, the spatial cognition debate is framed by developmentalist perspective. Cross-linguistic and cross-cultural comparison of the spatial term systems in adults is highly integrated with the research aiming at discovering some early developmental interrelations between language and the conceptualization of space, and with search for deep, both domain-specific and domain-general processes accounting for these relations. The relativist-universalist debate is still vivid there, but it takes different forms and arguments than in the color domain (see for example Levinson, 1996; Li & Gleitman, 2002; Newcomb & Uttal, 2006).

But even if an appropriate level of the debate is reached, it will still be far from clear what color or space categorization can tell about the role of language in cognition in general. Developmental data show that color and space (as well as number and time) categories are very special ones. Time-course and processes involved in object naming, learning biological categories, acquisition of verbs referring to common activities, or even abstract events (like mental terms) seems to be different (relatively quick, employing fast-mapping mechanism, guided by perceptual biases, like shape bias in object naming, or conceptual stances like relying on function and history in artifact naming). Many of these processes could also be partly shaped by language- and culture-specific factors (see e.g. Imai & Gentner, 1997, for the shape bias; Gopnik, Choi, & Baumberger, 1996, for nouns vs verbs acquisition, or Medin & Atran, 1999, and Tarlowski, 2008, for biological categories). However, the mechanisms of linguistic influence on the conceptual system can be different both across as well as within all these areas, and none of these studies alone can tell for ever which part of the universalist – relativist argument is right. Color categorization studied outside of the developmental frame is even less suitable to this aim.

References

Berlin, B. & Kay, P. (1969). *Basic Color Terms: their Universality and Evolution*. Berkeley: University of California Press.

Bornstein, M.H. (1985) On the development of color naming in young children: Data and theory. *Brain and Language, 26,* 72-93.

Brown, R. & Lenneberg, E. (1954). A study in language and cognition. *Journal of Abnormal and Social Psychology, 491,* 454-462.

Carruthers, P. (in press). Language in cognition. In E. Margolis, R. Samuels, & S. Stich (Eds.), *The Oxford Handbook of Philosophy of Cognitive Science*. Oxford: Oxford University Press.

Davidoff, J., Roberson, D., & Davies, I.R.L. (1999). Colour categories in a stone-age tribe. *Nature, 398,* 203-204.

Diamond, A. (2006). The early development of executive functions. In E. Bialystok & F.I.M. Craig (Eds.), *Life-span cognition: the mechanisms of change* (pp. 70-95). Oxford: Oxford University Press.

Franklin, A. & Davies, I.R.L. (2004). New evidence for infant color categories. *British Journal of Developmental Psychology, 22,* 349–377

Franklin. A., Drivonikou, G.V., Bevis, L., Davies I.R.L., Kay, P, & Regier, T. (2008). Categorical perception of color is lateralized to the right hemisphere in infants, but to the left hemisphere in adults. *Proceedings of the National Academy of Sciences, 105,* 3221–3225.

Gilbert, A.L., Regier, T., Kay, P., & Ivry, R.B. (2006). Whorf hypothesis is supported in the right visual field but not the left. *Proceedings of the National Academy of Sciences, 103,* 489–494

Gopnik, A., Choi, S., & Baumberger, T. (1996). Cross-linguistic differences in semantic and cognitive development. *Cognitive Development, 11* (2), 197-227

Heider, E. & Olivier, D. (1972). The structure of the color space in naming and memory for two languages. *Cognitive Psychology, 3,* 337-354.

Imai, M. & Gentner, D. (1996). A cross-linguistic study on constraints on early word meaning: Linguistic influence vs. universal ontology. *Cognition, 62,* 169-200

Kay, P. & Regier, T. (2006). Language, thought, and color: Recent developments. *Trends in Cognitive Sciences, 10,* 51-54

Levinson, S.(1996). Frames of reference and Molyneux's question: cross-linguistic evidence. In P. Bloom, M. Peterson, L. Nadel, & M. Garrett (Eds.), *Language and Space* (pp. 109-169). Cambridge, MA: MIT Press.

Li, P. & Gleitman, L. (2002). Turning the tables: language and spatial reasoning. *Cognition, 83,* 265-294.

Medin, D. & Atran, S. (Eds.) (1999). *Folkbiology*. Cambridge, MA: MIT Press.

Munnich, E. & Landau, B. (2003). The effects of spatial language on spatial representation: setting some boundaries. In D. Gentner & S. Goldin-Meadow (Eds.), *Language in Mind* (pp. 113-155). Cambridge, MA: MIT Press.

Newcomb, N.S., Uttal, D.H. (2006). Whorf versus Socrates: Round 10. *Trends in Cognitive Science, 10,* 394-396.

Regier, T. & Kay, P. (2009). Language, thought, and color: Whorf was half right. *Trends in Cognitive Science, 13,* 439-446.

Roberson, D. (2005). Color categories are culturally diverse in cognition as well as in language. *Cross-Cultural Research, 39,* 56-71

Roberson, D., Davidoff, J., & Braisby, N. (1999) Similarity and categorization: Neuropsychological evidence for a dissociation in explicit categorization tasks. *Cognition, 71,* 1-42.

Roberson, D., Davidoff, J., & Davies, I.R.L. (2000). Color categories are not universal: Replications & new evidence from a stone-age culture. *Journal of Experimental Psychology: General, 129,* 369-398.

Roberson, D., Davidoff, J., Davies, I.R.L., & Shapiro, L R. (2004). The development of color categories in two languages: A longitudinal study. *Journal of Experimental Psychology: General, 133,* 554-571.

Rosch, E.H. (1972). Probabilities, sampling and ethnographic method: The case of Dani color names. *Man, 7,* 448-466.

Shepard, R. (2001) Perceptual–cognitive universals as reflections of the world. *Behavioral and Brain Sciences, 24,* 581–601

Soja, N.N. (1994) Young children's concept of color and its relation to the acquisition of color words. *Child Development, 65,* 918-937.

Steels, L. & Belpaeme, T. (2005) Coordinating perceptually grounded categories through language: A case study for colour. *Behavioral and Brain Sciences, 28,* 469–489.

Tarlowski, A. (2008). Horizontal variability in biological knowledge. In F. Labrell & G. Chasseigne (Eds.), *Aspects du développment conceptuel et langagier* (pp.113-132). Paris: Éditions Publibook.

Wilcox, T. (1999). Object individuation: Infants' use of shape, size, pattern, and color. *Cognition, 72,* 125-166.