Language effects on the conceptualization of hybrids

NIRA MASHAL, YESHAYAHU SHEN, KARINE JOSPE and DAVID GIL

Language and Cognition / Volume 6 / Issue 02 / June 2014, pp 217 - 241
DOI: 10.1017/langcog.2014.6, Published online: 28 March 2014

Language and Cognition **6** (2014), 217–241. doi:10.1017/langcog.2014.6

Language effects on the conceptualization of hybrids*

NIRA MASHAL

School of Education, Bar Ilan University, Ramat Gan, Israel Gonda Multidisciplinary Brain Research Center, Bar-Ilan University, Ramat Gan, Israel

YESHAYAHU SHEN

Program of Cognitive Studies of Language Use, Tel Aviv University, Tel Aviv, Israel

KARINE JOSPE

Department of Education and Psychology, The Open University, Israel

AND

DAVID GIL

Linguistics, Max Plank Institute for Evolutionary Anthropology, Leipzig, Germany

(*Received* 24 *May* 2012 – *Revised* 27 *March* 2013 – *Accepted* 30 *June* 2013 – *First published online* 28 *March* 2014)

ABSTRACT

The current study investigates the conceptual hierarchy of humans–animals–plants–non-animate objects by using novel hybrids. Three experiments were conducted. In Experiment 1, twenty-one participants were presented with a grammatically asymmetrical phrase, in which the two components are associated with different linguistic properties, (e.g., *a man with a horse's head*) followed by a visual hybrid, and were asked to judge whether the phrase described the hybrid. In Experiment 2, thirty participants were presented with a visual hybrid and were asked to categorize it according to one of its visually presented components in a forced-choice judgment task. In Experiment 3, twenty-nine participants were presented with a visual hybrid that followed a grammatically

[*] This research was supported by Grant No. 969-07 from the Israel Science Foundation of the Israel Academy of Sciences and Humanities to the second author. Address for correspondence: Dr Nira Mashal, School of Education, Bar-Ilan University, Israel. tel: (972)-3-5317178; fax: (972)-3-7384029; e-mail: nmashal2@gmail.com

symmetrical phrase, in which both components carry similar grammatical properties (e.g., *half-human half-horse*), and were asked to judge whether the phrase described the hybrid. A conceptual hierarchy effect was found in Experiment 1 but not in the other two experiments. These findings show that the hierarchy effect occurs only in verbal tasks that involve asymmetrical grammatical constructions. We suggest that the pragmatic tendency to map the hierarchically higher concept onto the higher grammatical function applies to asymmetrical constructions but not to symmetrical constructions.

KEYWORDS: conceptual hierarchy, hybrids, categorization

1. Introduction

A visual hybrid is a novel image that combines components of two or more familiar objects belonging to different ontological categories. Well-known examples of such hybrids are mythological creatures such as the half-human half-horse centaur or the half-human half-fish mermaid (Carroll, 1994; Forceville, 1996). Hybrid conceptualization involves a basic form of knowledge organization, called the Conceptual Hierarchy (CH), in which concepts such as humans, animals, plants, and non-animates are hierarchically organized (e.g., Connor & Kogan, 1980; Deane, 1992; Keil, 1979). Here we focus on four levels of the hierarchy which are relevant for the creation of visual hybrids: HUMANS – ANIMALS – PLANTS – NON-ANIMATE. The CH reflects the relative prominence of each ontological category, with humans being the most prominent or the most salient conceptual category and non-animate being the least prominent category (Deane, 1992; Levin & Hovav, 2005). Other levels of the hierarchy, such as 'events' or 'abstract concepts', cannot be used, at least not in a straightforward way, to create visual hybrids, as such hybrids consist of concrete and visualizable components.

1.1. THE CONCEPTUAL HIERARCHY (CH)

The CH has been shown to play a central role in a variety of psychological and linguistic phenomena. For instance, Keil (1981) argued that a very similar hierarchy emerges in cognitive development, with concepts placed at the top of the hierarchy emerging first and concepts placed at lower levels emerging later. Furthermore, the Animacy Hierarchy (see Deane, 1992, for a review of various versions of this hierarchy) influences various linguistic phenomena such as word order, pronominalization patterns, and assignment of grammatical case (Deane, 1992). The Animacy Hierarchy has also been shown to be central to language production in that language users tend to encode nouns that are higher on the hierarchy in 'higher' grammatical functions (Branigan, Pickering, & Tanaka, 2008; Feleki, 1996). We will return to the latter point below.

Other studies have examined the CH in more creative conceptual combinations. For example, Kogan and his colleagues (Connor & Kogan, 1980; Kogan, Mindi, & Heleen, 1989) studied the CH in metaphors (see also Deane, 1993), with the Metaphorical Triad Task (MTT). On this task, participants are asked to generate a verbal description of metaphorically related images (e.g., an image of an old, bent, and tired man and a barren, gnarled tree) in the 'A is like B' form. It was found that items high on the conceptual hierarchy (e.g., the old man) were more likely to function as the subject or topic of the comparison, and items low on the hierarchy (e.g., the gnarled tree) were more likely to function as the vehicle or source.

These studies suggest that processing of various verbal and pictorial stimuli is sensitive to the conceptual hierarchy. Our main question is whether the conceptualization of novel conceptual combinations, such as visual hybrids, is also sensitive to the CH.

1.2. VISUAL HYBRIDS AND THE CONCEPTUAL HIERARCHY

Hybrids differ radically from the type of stimuli used in previous studies (notably Connor & Kogan, 1980; Kogan et al., 1989) in that they do not map onto pre-stored conceptual representations as do other verbal constructions. However, the fact that hybrids combine concepts that belong to different ontological categories raises the possibility that the CH will affect their processing. The question, then, is whether the relative prominence of the two components that constitute the hybrid is determined by the CH. For example, the question is whether in the 'half-man–half-horse' hybrid, the human will be more prominent than the animal.

In a previous study, Shen and Gil (2010) examined the extent to which hybrid conceptualization is sensitive to CH. Specifically, they asked whether individuals who are asked to categorize hybrids do so on the basis of the CH. Shen and Gil (2010) focused on four levels within this hierarchy, including humans, animals, plants, and physical objects, in this order. Shen and Gil's study yielded a surprising difference between the verbal and visual conceptualization of hybrids. A CH effect was found on a task in which participants were asked to produce a verbal description to a visual hybrid (e.g., the half-human half-bird in Figure 1A). Participants preferred to describe the hybrid as *a man with bird wings*, rather than as *a bird with human legs*. That is, they preferred to categorize the hybrid as a human being rather than to describe it as a bird (regardless of top-down orientation). Similarly, in a forced-choice experiment, participants preferred the former description

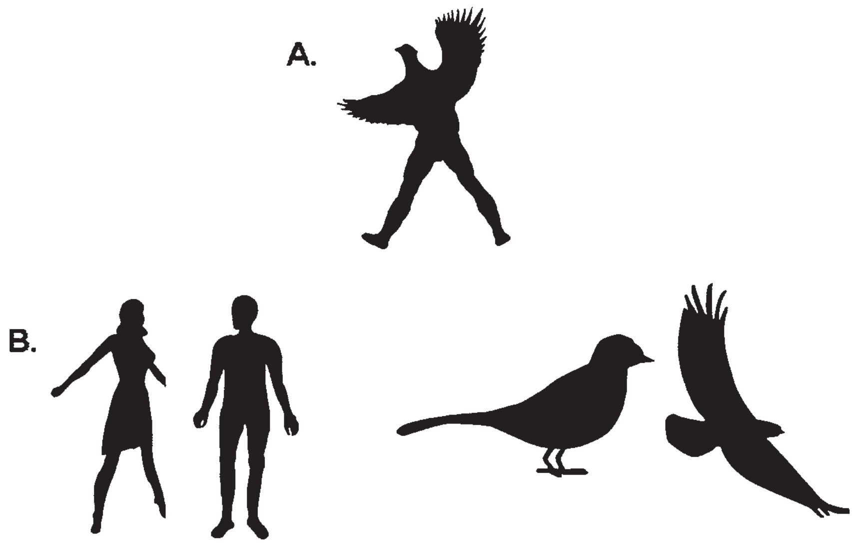

Fig. 1. (A) An example of a visual hybrid (half-human half-bird); (B) Visually presented categories (humans, birds) that match one of the components of the corresponding hybrid.

to the latter one. In contrast, when non-verbal (visual) tasks were used, no such effect was found. Instead, on a non-verbal task, in which participants were asked to categorize a given hybrid by choosing one of two visually presented categories, conceptual hierarchy had no effect. For example, participants were presented with a half-human half-bird hybrid (Figure 1A) and were asked to determine if it belonged to humans or to birds (Figure 1B). Categorization was at chance level, with participants demonstrating no preference for either category. A similar pattern was found on other tasks that involved no language.

Note, however, that it is possible that hybrids are not categorized in terms of the CH because they do not map onto familiar concepts. Although the objects that constitute the hybrid are conceptually represented, they differ in terms of their hierarchy, and thus no categorization preference emerges. This interpretation might imply that the CH does not apply to hybrid categorization. As Wagner, Kronberger, Nagata, Sen, Holtz, & Palacios, (2010, p. 235) state: "besides being unfamiliar, [people] will perceive the hybrid as highly unnatural, unpleasant, negative, as a threat to the received 'natural order' and as lacking a category identity. Hybrids transcend the categorical grid of the local world … and imply category confusion." This category confusion may account for the fact that there was no CH effect in our previously administered visual categorization task.

Nevertheless, when asked to describe a hybrid verbally, people must convert their conceptualization into a grammatical construction, typically a sentence or a noun phrase. It has been established by both linguistic (e.g., Deane, 1992; Levin & Hovav, 2005) and psycholinguistic (e.g., Branigan et al., 2008; Feleki, 1996) studies that in grammatical constructions that involve subject–predicate in a sentence or head–modifier in a noun phrase, the higher grammatical function (the subject or the head noun) usually encodes the hierarchically higher concept. For example, the sentence *The man is walking with the dog* sounds more natural than the sentence *The dog is walking with the man* (see also Shen & Gil, 2010). Thus, participants tend to conform to the CH when generating a verbal description or when judging which of two (verbal) descriptions is more adequate for a given hybrid. Accordingly, a description such as *a man with bird wings* will be preferred to a description such as *a bird with human legs*. There is some evidence that the pragmatic tendency to encode nouns that are higher on the conceptual hierarchy in higher grammatical functions is applied to asymmetrical constructions (Branigan et al., 2008; Feleki, 1996; Shen & Gil, 2010).

Alternatively, the CH preference may reflect the actual hybrid categorization process, not only the selection of verbal descriptions. Indeed, priming a visual categorization task through verbal descriptions increased the CH effect (Shen, Gil, & Roman, 2006). That is, when participants were asked to categorize a visual hybrid as either human or animal, they categorized it as human more often if it was primed by the phrase *a man with bird wings* as compared with cases in which the verbal description was incompatible with the CH (e.g., *a bird with human le*gs).

This 'language effect' in hybrid conceptualization refers to the long-standing and intensive debate as to whether language shapes conceptual, non-verbal, processes (e.g., Gentner & Goldin-Meadow, 2003; Gleitman & Papafragou, 2005; Lupyan, 2012; Wolff & Holmes, 2011).

1.3. LANGUAGE EFFECT ON CATEGORIZATION

Past research examined language effects on perceptual processing as well as its effect on higher cognitive processes such as relational thinking (Loewenstein & Gentner, 2005), false belief (de Villiers & de Villiers, 2003), and categorization (Wolff & Holmes, 2011). Language was found to affect various aspects of categorization processes. For example, Waxman and Markow (1995) maintained that language may help forming new categories so that infants learn the noun–category linkage at the very beginning of lexical acquisition, thus showing an initial bias to interpret a word applied to an object as referring to that object and to other members of its kind. Lupyan, Rakison, and McClelland (2007) found that the very use of (nonsensical) verbal labels affects categorization, in that it facilitates category learning of things to approach versus things to avoid. Other studies demonstrated the effect of verbal labels on color categorization (Roberson, Davidoff, Davies, & Shapiro, 2005), as well as the effect of gender marking on gender categorization (Boroditsky, Schmidt, & Phillips, 2003).

The influence that conceptual hierarchy might have on categorization, and how this influence is related to language use, has not received much attention in the literature. The pragmatic tendency to place more prominent ontological categories (e.g., animates vs. non-animates, or humans vs. non-humans) in higher grammatical functions (subject vs. object, head vs. modifier) might mean that linguistic labeling will affect hybrid processing as well. Hence, our study focuses on the ways in which language might enhance the hierarchical conceptualization of visual hybrids.

1.4. THE PRESENT STUDY

Shen et al. (2006) have provided some initial evidence for the existence of an effect of language on the processing of conceptual hierarchy with non-verbal stimuli. The present study aims to extend our previous work by addressing two major issues that have not yet been addressed. First, our previous study detected a difference between visual and verbal presentation through an off-line task. Here we wanted to examine whether this difference reflects the outcome of a delayed, elaborative process or whether it reflects a more basic process that could also be documented in on-line tasks. Experiment 1 was designed to test whether participants would be faster to match a hybrid image to a verbal description compatible with the hierarchy (e.g., *a man with bird wings*) than to match a visual hybrid to a verbal description that was incompatible with the hierarchy (e.g., *a bird with a human head*). Experiment 2 was designed to examine whether the CH effect is restricted to verbal stimuli or appears in non-verbal stimuli as well.

Experiment 3 was designed to examine whether the hierarchy effect would be found not only in grammatically asymmetrical constructions such as *a human with bird wings*, but also in grammatically symmetrical constructions such as *half-human half-bird*.

Grammatically symmetrical structures are ones in which both components carry similar grammatical and semantic properties, as in coordination (e.g., 'half-human half-bird'). In contrast, grammatically asymmetrical structures are ones in which the two components are associated with different linguistic properties. In these constructions, constituents may exhibit different morphological marking, they may form a head–modifier relationship, or they may belong to different semantic categories. A typical example of an asymmetrical structure is the genitive or attributive possessive (e.g., 'a human with bird wings').

Note that within the phrase *a human with bird wings* the head and the modifier have asymmetrical grammatical roles, whereas noun phrase conjunctions such as *half-human half-bird* are grammatically symmetrical. It is possible that individuals show preference to human-initial phrases because in asymmetrical phrases the conceptual hierarchy matches grammatical roles. Manipulating the order of nouns within the symmetrical phrases can shed light on this possibility. Hence, in Experiment 3 we asked participants to match a hybrid image to either a hierarchy-compatible grammatically symmetrical verbal description (e.g., *half-human half-bird*) or to a hierarchy-incompatible description (e.g., *half-bird half-human*).

It is important to note that Tanaka, Holly, and Martin (2005) and Branigan et al. (2008) have recently suggested that CH does not affect the ordering of nouns in noun phrase conjunctions (i.e., in grammatically symmetrical constructions). In contrast, according to Cooper and Ross (1975), the CH does affect noun order in a conjunction construction. We assume that language plays a more dominant role in the processing of grammatically asymmetrical constructions than in the processing of symmetrical phrases, because in the asymmetrical constructions both the order of nouns and the grammatical structure match the CH. Hence, we expect to find no CH effect in grammatically symmetrical phrases.

2. Experiment 1: investigating conceptual hierarchy effects in a verbal task with asymmetrical grammatical phrases

The goal of this experiment was to examine whether a grammatically asymmetrical hierarchically compatible verbal description (e.g., *a man with bird wings*) would lead to a faster and more accurate response to a hybrid image relative to a hierarchically incompatible verbal description (e.g., *a bird with human head*).[1]

2.1. METHOD

2.1.1. *Participants*

Twenty-one Bar Ilan University undergraduate students (5 men), aged 20–36 (mean age = 26.7, *SD* = 3.5), participated in the experiment for course credit. All participants were native Hebrew speakers and right-handed.

[1] The experiment was conducted in Hebrew. A number of studies have addressed the issue of animacy hierarchy in Hebrew (see Merchant, 2006). As far as the present study is concerned, Hebrew exhibits a similar use of animacy properties as do corresponding languages such as English.

TABLE 1. *Examples of the stimuli used in Experiments* 1 *and* 3

	Phrase	**Type of phrase**	**Hybrid image**	**Condition**	**Hybrid spatial orientation**
1	*A man with a lizard tail Half-human half-lizard*	Hierarchically-compatible		Congruent hierarchically-compatible	Top part is higher on the CH
2	*A cactus with human legs Half-cactus half-human*	Hierarchically-incompatible		Congruent hierarchically-incompatible	Top part is lower on the CH
3	*A hammer with a hoe Half-hammer half-hoe*	Neutral		Congruent neutral	Both parts belongs to the same CH level
4	*An elephant with a boat Half-elephant half-boat*	Hierarchically-compatible		Incongruent hierarchically-compatible	Top part is high in the hierarchy
	A flower with a man Half-flower half-human	Hierarchically-incompatible		Incongruent hierarchically-incompatible	Top part is low in the hierarchy
5	*A boat with a human Half-boat half-human*	Hierarchically-compatible or incompatible		Incongruent neutral	Both parts belongs to the same hierarchic level

2.1.2. *Stimuli construction*

The stimulus pool consisted of eighty hybrid images that included forty true hybrids (i.e., hybrids whose components belong to two different categories) and forty neutral hybrids (i.e., hybrids whose components belong to the same ontological category). All visual images were designed by a graphic artist using Adobe FreeHand software. Four ontological categories (human, animal, plant, and object) were used to create true hybrids that included six possible types of hybrid: human–animal, human–plant, human–object, animal–plant, animal–object, and plant–object. Each component represents an easily recognizable object of a familiar item (e.g., a tree trunk that is easily recognizable as part of a tree). Each type of combination included eight different hybrids (a total of 6 × 8 = 48 images). The images were controlled for spatial orientation, with objects higher on the CH constituting the top part of the hybrids in half the stimuli, and the bottom part in the other half. For example, for the human–animal combination, the human was placed on the top and the animal at the bottom in four (out of eight) hybrids, while the animal was placed at the top and the human at the bottom in the remaining four hybrids. Images were controlled for their relative size, so that the two components were roughly the same size. All hybrids were black images on a white background and were vertically oriented. The forty neutral hybrids were created by combining two objects from the same ontological category (e.g., *half-lion half-dog*).

Each visual image appeared after a verbal phrase that could be either hierarchically compatible or hierarchically incompatible. Hierarchically compatible phrases corresponded with the CH (e.g., *a man with a lizard tail*). Hierarchically incompatible phrases violated the CH (e.g., *a tree with elephant legs*). Phrases were in the form X *with* Y (the phrase has the same grammatical structure in Hebrew: X *im* Y).

The combination of visual images and verbal phrases resulted in the following conditions (see Table 1 for examples):

> Congruent hierarchically-compatible condition: twenty true visual hybrids appeared after hierarchically-compatible phrases that matched the hybrids (e.g., a visual image of *half-man half-lizard* appeared after the phrase *a man with a lizard tail*).
> Congruent hierarchically-incompatible condition: twenty true visual hybrids appeared after hierarchically-incompatible phrases that matched the hybrids (e.g., the visual image *half-cactus half-man* appeared after the phrase *a cactus with human legs*).
> Congruent neutral condition: twenty neutral images appeared after phrases that matched the images (e.g., an image of *half-lion half-dog* appeared after the phrase *a lion with a dog*).
> Incongruent neutral condition: twenty neutral images appeared after phrases that did not describe the images (e.g., an image of *half-lion half-dog* appeared after the phrase *a hammer with a hoe*).

Each stimulus was comprised of a phrase followed by a hybrid image. For each hybrid image two phrases are presented. The upper phrase (*an* X *with a* Y) was used in Experiment 1, and the lower phrase (*half*-X *half*-Y) was used in the Experiment 3. Phrases are literally translated from Hebrew.

Incongruent hierarchically-compatible condition: twenty true visual hybrids appeared after hierarchically-compatible phrases that did not describe the images (e.g., an image of *half-bird half-scissors* appeared after the phrase *a hammer with a hoe*).

Incongruent hierarchically-incompatible condition: twenty true visual hybrids followed hierarchically-incompatible phrases that did not describe the images (e.g., an image of *half a hammer half a plant* appeared after the phrase *a flower with a man*).

In summary, the experiment consisted of six conditions: two congruent conditions of true visual hybrids, one congruent condition of neutral images, two incongruent conditions of true visual hybrids, and one incongruent neutral condition. Thus, half the stimuli appeared in congruent conditions and half in incongruent conditions.

2.1.3. *Procedure*

Participants were presented with a verbal phrase and then a visual image. The phrase appeared at the center of the screen for 1800 ms, then the screen was blank for 300 ms, and then the hybrid appeared for 1000 ms. Once the hybrid disappeared, a fixation point appeared for 2000 ms and then another trial began. Participants were instructed to indicate whether the phrase described the following image by pressing with the right index finger the right (key 'n') or left (key 'b') of the keyboard. We expected a positive response for all congruent conditions and a negative response for all incongruent conditions. Assignment of the keys to *yes* or *no* responses was counterbalanced across participants. The experiment was run by SuperLab software (version 4.0; Abboud, 1991). Stimulus presentation was divided into two blocks, randomly ordered across participants. There were three practice trials before the experiment.

2.2. RESULTS

Analyses were performed on reaction times (RT) to correct responses and on accuracy level (percent of correct responses). Responses that were slower than two standard deviations from the mean RT of each participant were removed from the analysis.

2.2.1. *RT analysis*

A one-way repeated measures analysis of variance (ANOVA) of RT on the incongruent hierarchically-compatible condition and the incongruent

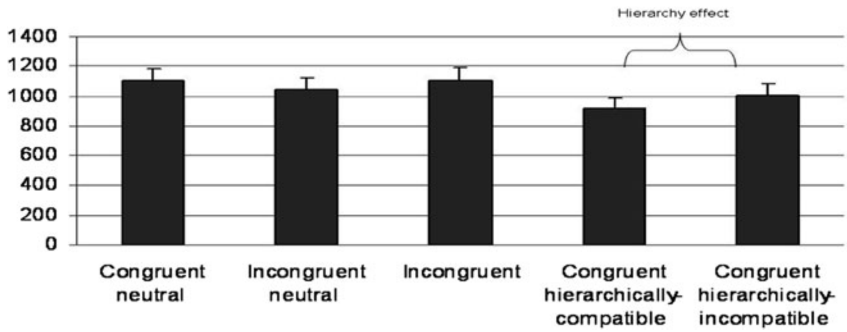

Fig. 2. Mean reaction times (ms) in Experiment 1.

hierarchically-incompatible condition revealed no statistically significant difference ($F(1,20) = -0.94$, n.s.). Following this analysis, mean RT was recalculated to include both incongruent conditions together. All further analyses were conducted on five conditions: Three congruent conditions (two with true visual hybrids, one with neutral images), and two incongruent conditions (one with true visual hybrids, and one with neutral images).

A one-way repeated measures ANOVA with five conditions was performed. There was a significant main effect of condition ($F(4,80) = 11.17$, $p < .0001$. A Tukey HSD post-hoc analysis revealed a hierarchy effect within the first two congruent conditions: participants responded more quickly to true hybrids that followed hierarchically-compatible phrases (920 ms) than to true hybrids that followed hierarchically-incompatible phrases (1008 ms) ($p < .05$) (see Figure 2 and Table 2 for mean RT).

A Tukey HSD post-hoc analysis revealed that RT on the congruent hierarchically-compatible hybrid condition (920 ms) were shorter than were RT on the incongruent hybrid condition (1102 ms) ($p < .001$). RT on the congruent hierarchically-incompatible hybrid condition (1008 ms) were shorter than were RT on the incongruent hybrid condition (1102 ms) ($p < .05$). Also, participants responded more quickly on the congruent hierarchically-compatible hybrid condition (920 ms) than on the congruent neutral condition (1100 ms) ($p < .01$).

A repeated measures ANOVA was conducted with RT as the dependent variable, hierarchical compatibility condition as the within-subject factor, and orientation (top, down) as the between-subject factor. As expected, the main effect of condition was significant ($F(1,52) = 6.63$, $p < .05$), indicating that participants responded more quickly to true hybrids that followed hierarchically-compatible phrases (920 ms) than to true hybrids that followed hierarchically-incompatible phrases (1008 ms) ($p < .05$). No significant difference in RT was found between hybrids in which the top part was higher on the CH

TABLE 2. *Mean reaction time (ms), percent of correct responses, and significant results in Experiment* 1

	Congruent hierarchically-compatible	Congruent hierarchically-incompatible	Congruent neutral	Incongruent	Incongruent neutral
RT (ms)	920	1008	1100	1102	1041
SD	305	360	385	419	395
Accuracy (%)	97.62	93.85	94.64	96.73	95.04
SD	3.63	5.84	6.20	3.46	8.19

(962 ms) and hybrids in which the bottom part was higher on the CH (967 ms) (n.s.). The condition × orientation interaction was not significant ($F(1,52) = 1.62$, n.s.).

2.2.2. *Accuracy analysis*

A one-way repeated measures ANOVA of the mean percent of correct responses across the five conditions was performed. There was a significant main effect of condition ($F(4,80) = 18.26$, $p < .05$). A Tukey HSD post-hoc analysis revealed a CH effect within the first two congruent conditions: participants responded more accurately to true hybrids that followed hierarchically-compatible phrases (97.62%) than to true hybrids that followed hierarchically-incompatible phrases (93.85%) ($p < .05$) (see Table 2 for mean percent of correct responses). No significant difference in accuracy was found between the other conditions.

Given that our design measured both RT and accuracy, we looked for a speed–accuracy trade-off. We thus compared the accuracy rate of the fastest (participants with RTs above the average) and slowest (participants with RTs below the average) participants in the hierarchically-compatible condition and found no significant differences between them ($t(19) = 0.21$, p = n.s.). We then compared the accuracy rate of the fastest and slowest participants in the hierarchically-incompatible condition and found no significant differences between them ($t(19) = 0.89$, p = n.s.).

To examine whether it was easier to match hybrids according to a specific combination of categories relative to the other combinations, we performed a one-way repeated measures ANOVA with type of combination (human–animal, human–plant, human–object, animal–plant, animal–object, plant–object) as the independent variable and RT as the dependent variable. Mean RT for correct responses is shown in Table 3. The analysis showed a significant difference among the six combinations ($F(5,95) = 3.99$, $p < .05$). Tukey HSD post-hoc comparison revealed that responses to the animal–plant hybrids

TABLE 3. *Mean reaction times and percent of correct responses in Experiment* 1

	Human–animal	Human–plant	Human–object	Animal–plant	Animal–object	Plant–object
RT (ms)	900	908	888	1014	917	920
SD	289	309	268	277	278	292
Accuracy (%)	92.50	96.88	91.25	89.38	96.25	97.50
SD	11.03	5.55	18.63	11.67	9.16	5.13

were significantly slower than were responses to the human–animal hybrids ($p < .01$), significantly slower than were responses to the human–plant hybrids ($p < .05$), significantly slower than were responses to the human–objects hybrids ($p < .01$), significantly slower than were responses to the animal–object hybrids ($p < .05$), and significantly slower than were responses to the plant–object hybrids ($p < .05$). No other comparisons were significant. In addition, the percent of correct responses did not differ across the six categories ($F(5,95) = 1.94$, n.s.).

2.3. MIXED ANALYSIS

We performed a mixed-model analysis with subjects and items as random effects using SPSS 15.0 software for Windows (SPSS Inc., Chicago, IL) in order to test whether the effects that appear significant by subjects remain significant by more appropriate analyses that treat subjects and items as simultaneous random factors. The mixed-model analysis has a lower risk of Type I error (Quene & van den Bergh, 2008). The dependent variable (RT) was not normally distributed and thus the analysis was performed on log (RT) of correct responses. Compound symmetry was assumed between items. Responses below two standard deviations from the mean RT of each participant were removed from the analysis. There was a significant condition effect ($F(4,2134) = 23.33$, $p < .0001$). Results revealed a hierarchy effect within the first two congruent conditions. Thus, participants responded more quickly to true hybrids that followed hierarchically-compatible phrases (920 ms) than to true hybrids that followed hierarchically-incompatible phrases (1008 ms) ($t(2134) = -2.46$, $p < .05$). Additional results indicated that reaction times in the congruent hierarchically-incompatible condition (1008 ms) were shorter than reaction times in the incongruent hybrid condition (1102 ms) ($t(2134) = 4.91$, $p < .001$). Also, participants responded more quickly in the congruent hierarchically-incompatible hybrid condition (1008 ms) than in the congruent neutral condition (1100 ms) ($t(2134) = 5.64$, $p < .001$). Thus, as expected, our results documented a hierarchical effect for hybrids that follow asymmetrical grammatical constructions.

3. Experiment 2: investigating conceptual hierarchy effects in a visual categorization task

In the first experiment we observed a hierarchy effect with verbal phrases, so that participants matched hybrids to phrases more quickly and more accurately when phrases matched the CH than when phrases did not match the CH. The goal of the second experiment was to investigate whether the same pattern of responses is found in a non-verbal task. That is, the main question is whether the CH effect reflects the processing of the hybrid itself or the processing of the accompanying verbal phrase. We used a non-verbal forced-choice categorization task to examine this question.

3.1. METHOD

3.1.1. *Participants*

Thirty Bar Ilan undergraduate students (14 men), aged 19–29 (mean age = 24.3, *SD* = 3.2), participated in the study for course credit. All participants were native Hebrew speakers and right-handed. None of the participants in Experiment 2 participated in Experiment 1.

3.1.2. *Stimuli construction*

The stimulus pool consisted of ninety-six hybrids, belonging to two congruent conditions and one incongruent condition (see Table 4). The congruent conditions included forty-eight hybrids that appeared prior to an image that represented the category of one of the components of the corresponding visual image. The congruent hierarchically-compatible condition included categories that represented the hierarchically-higher part of the hybrid (for example, a *half-man half-lizard* hybrid appeared prior to an image of a group of three people). The congruent hierarchically-incompatible condition included categories that represented the hierarchically-lower part of the hybrid (for example, a *half-bird half-plant* hybrid followed by an image of a group of three different plants).

Forty-eight hybrids were assigned to an incongruent condition in which the category image did not represent the category of either component of the corresponding hybrid stimuli. For example, a *half-kettle half-plant* hybrid was followed by an image of a group of three fish. To control for spatial orientation, the category image represented the top part of the hybrid in half the cases, and the bottom part in the other half.

3.1.3. *Procedure*

Participants were presented with a visual image of a hybrid and then an image of a group. The hybrid appeared at the center of the screen for 2000 ms. Once

TABLE 4. *Examples of the stimuli used in Experiment* 2

	Hybrid	Category	Condition	Category spatial orientation
1			Congruent hierarchically-compatible	Category image refers to the top part of the hybrid
2			Congruent hierarchically-incompatible	Category image refers to the bottom part of the hybrid
3			Incongruent	Category image does not refer to the hybrid

the hybrid disappeared, the screen was blank for 300 ms, and then a category image appeared for 1000 ms. Next, a fixation point appeared at the center of the screen for 1500 ms and a new trial began. Participants were instructed to indicate as rapidly as possible whether the hybrid was a member of the category depicted by the second image by pressing with the right index finger the right (key 'n') or left (key 'b') of the keyboard. Assignment of the keys to *yes* or *no* responses was counterbalanced across participants. The experiment was run by SuperLab software (version 4.0; Abboud, 1991). The experiment began with three trials of hybrid–category practice.

3.2. RESULTS

Analyses were performed for RT to correct responses and for accuracy rate (percent of correct responses). Responses below two standard deviations from the mean RT of each participant were removed from the analysis.

3.2.1. *RT analysis*

A two-tailed paired samples t-test that compared responses on the two congruent conditions revealed no significant difference in RT between the congruent hierarchically-compatible condition (1042 ms) and the congruent hierarchically-incompatible condition (1045 ms) ($t(29) = -0.098$, n.s.).

In order to examine the effect of spatial orientation, a repeated measures ANOVA was conducted with RT as the dependent variable, hierarchical compatibility as the within-subject factor, and orientation (top, down) as the between-subject factor. The main effect of hierarchical compatibility was not significant ($F(1,58) = 0.21$, n.s.). The effect of orientation was not significant either (mean RT when top component higher on CH = 1033 ms, mean RT when bottom component higher on CH = 1065) ($F(1,58) = 0.22$, n.s.). The condition-orientation interaction was significant ($F(1,58) = 24.92$, $p < .001$). Tukey post-hoc analysis revealed that participants responded more quickly to hierarchically-compatible congruent hybrids in which the top part was higher on the CH (988 ms) than to hybrids in which the bottom part was higher on the CH (1119 ms) ($p < .01$).

3.2.2. *Accuracy*

A two-tailed paired samples *t*-test that compared the accuracy rate on the two congruent conditions revealed no significant difference between the hierarchically-compatible condition (74.58%) and the hierarchically-incompatible condition (78.75%) ($t(29) = -1.89$, n.s.). Participants made an average of 23.3% errors in judging hybrids on the congruent conditions, and 10.7% errors on the incongruent conditions, and the difference between error rates was statistically significant ($t(29) = -2.74$, $p < .05$). This finding is not surprising, given that the incongruent condition involved categorical images that represented neither of the two components of the hybrid.

Because our design measured both RT and accuracy, we looked for a speed–accuracy trade-off. We thus compared the accuracy rate of the fastest in the hierarchically-compatible condition (participants with RTs above the average) and slowest participants (participants with RTs below the average) and found no significant differences between them ($t(28) = 1.54$, p = n.s.). We then compared the accuracy rate of the fastest in the hierarchically-incompatible condition (participants with RTs above the average) and slowest participants (participants with RTs below the average) and found no significant differences between them ($t(28) = 1.79$, p = n.s.).

To examine whether it was easier to categorize hybrids according to a specific category relative to the other categories, we performed a one-way repeated measures ANOVA with type of category (humans, animals, plants, and objects) as the independent variable and RT as the dependent variable. Mean RT for correct responses is shown in Table 5. The analysis showed no significant difference among the four categories ($F(3,87) = 0.57$, n.s.).

TABLE 5. *Mean reaction time and percent of correct responses in Experiment* 2

	Human	Animal	Plant	Object
RT (ms)	1071	1068	1080	1098
SD	336	228	317	266
Accuracy (%)	71.21	77.18	80	77.78
SD	27.38	22.95	19.67	23.01

In addition, the percent of correct responses did not differ across the four categories ($F(3,87) = 1.34$, n.s.).

Thus, the results of Experiment 2 show no CH effect on a non-verbal categorization task.

3.2.3. *Mixed analysis*

We performed a mixed-model analysis with subjects and items as random effects using SPSS 15.0 software for Windows (SPSS Inc., Chicago, IL) in order to test whether the effects that appear significant by subjects remain significant by more appropriate analyses that treat subjects and items as simultaneous random factors. The dependent variable (RT) was not normally distributed, therefore analysis was performed on log (RT) of correct responses. Compound symmetry was assumed between items. Responses below two standard deviations from the mean RT of each participant were removed from the analysis. There was no significant condition effect ($F(1,2363) = 3.27$, n.s.). Thus, results revealed no significant difference in RT between the congruent hierarchically-compatible condition (1042 ms) and the congruent hierarchically-incompatible condition (1045 ms) ($t(2363) = -1.80$, n.s.).

Thus, the results of both analyses of Experiment 2 revealed no significant differences between the congruent hierarchically-compatible condition and the congruent hierarchically-incompatible condition.

4. Experiment 3: investigating conceptual hierarchy effects in a verbal task with grammatically symmetrical phrases

The goal of this experiment was to investigate whether people match a hybrid image to a hierarchically compatible, grammatically symmetrical verbal description (e.g., *half-human half-bird*) faster and more accurately than they match a hybrid image to the corresponding hierarchically incompatible description (e.g., *half-bird half-human*). We also wanted to see whether grammatically symmetrical phrases lead to the same CH effect seen in grammatically asymmetrical phrases.

4.1. METHOD

4.1.1. *Participants*

Twenty-nine Bar Ilan undergraduate students (9 men), aged 20–33 (mean age = 27.4, *SD* = 4.1), participated in the study for course credit. All participants were native Hebrew speakers and right-handed. None of the participants in Experiment 3 participated in any of the other two experiments.

4.1.2. *Stimuli*

We used the same stimuli from Experiment 1, but changed the verbal phrases from *an* X *with a* Y into *half*-X *half*-Y.

4.1.3. *Procedure*

The procedure was identical to the procedure of Experiment 1.

4.2. RESULTS

Data from two participants were not analyzed due to technical problems during their testing (i.e., unexpected noise). Analyses were performed for RT to correct responses and for accuracy rate (percent of correct responses). Responses that were below two standard deviations from the mean RT of each participant were removed from the analysis.

4.2.1. *RT analysis*

A one-way repeated measures ANOVA with five conditions (three congruent and two incongruent conditions) was performed. There was a significant main effect of condition ($F(4,104) = 8.43$, $p < .01$) (see Table 6 for mean RT in each condition). A Tukey HSD post-hoc analysis did not find a hierarchy effect, so that there was no significant difference between responses on the congruent hierarchically-compatible hybrid condition (923 ms) and responses on the congruent hierarchically-incompatible hybrid condition (951 ms) (n.s.). Additional analyses showed that participants responded more quickly on the congruent hierarchically-compatible condition (923 ms) than on the incongruent condition (1002 ms) ($p < .01$). Reaction times on the congruent hierarchically-compatible hybrid condition (923 ms) were shorter than were reaction times on the congruent neutral condition (1030 ms) ($p < .001$).

TABLE 6. *Mean RT and percent of correct responses in Experiment* 3

	Human–animal	Human–plant	Human–object	Animal–plant	Animal–object	Plant–object
RT (ms)	907	920	943	952	951	944
SD	323	323	305	361	361	320
Accuracy (%)	98.62	98.18	99.08	96.43	98.78	96.86
SD	3.93	4.44	4.72	6.26	6.30	7.29

Reaction times on the congruent hierarchically-incompatible condition (952 ms) were shorter than were reaction time on the congruent neutral condition (1030 ms) ($p < .01$).

A repeated measures ANOVA was conducted with RT as the dependent variable, hierarchical compatibility condition as the within-subject factor, and orientation (top, down) as the between-subject factor. The main effect of condition was not significant ($F(1,38) = 0.01$, n.s.). The main effect of orientation was also not significant, indicating no difference in RT between hybrids in which the top part was higher on the CH (934 ms) and hybrids in which the bottom part was higher on the CH (932 ms) ($F(1,38) = 0.97$, n.s.). The condition × orientation interaction was also not significant ($F(1,38) = 0.08$, n.s.).

4.2.2. *Accuracy analysis*

A one-way repeated measures ANOVA with five conditions (three congruent and two incongruent conditions) was performed on the percent of correct responses. There was a significant main effect of condition ($F(4,104) = 3.90$, $p < .01$). A Tukey HSD post-hoc analysis found no hierarchy effect, so that participants responded equally accurately on the congruent hierarchically-compatible condition (98.45%) and on the congruent hierarchically-incompatible condition (97.41%) (n.s.). Another Tukey HSD post-hoc analysis revealed that the accuracy rate on the congruent hierarchically-compatible condition (98.45%) was higher than was the accuracy rate on the congruent neutral condition (95.34%) ($p < .05$). The accuracy rate on the congruent hierarchically-compatible condition (98.45%) was also higher than was the accuracy rate on the incongruent condition (94.88%) ($p < .01$).

Moreover, we looked for a speed–accuracy trade-off. We thus compared the accuracy rate of the fastest in the hierarchically-compatible condition (participants with RTs above the average) and slowest participants (participants with RTs below the average) and found no significant differences between them ($t(27) = 9.76$, $p =$ n.s.). We then compared the accuracy rate of

TABLE 7. *Mean RT and percent of correct responses in Experiment* 3

	Congruent hierarchically-compatible	Congruent hierarchically-incompatible	Congruent neutral	Incongruent	Incongruent neutral
RT	923	951	1030	1002	966
SD	293.30	307.48	326.27	284.46	292.06
Accuracy	98.45	97.41	95.34	94.88	97.24
SD	3.24	4.07	5.07	6.47	6.22

the fastest in the hierarchically-incompatible condition (participants with RTs above the average) and slowest participants (participants with RTs below the average) and found no significant differences between them ($t(27) = 0.56$, p = n.s.).

To examine whether it was easier to match hybrids according to a specific combination of categories relative to the other combinations, we performed a one-way repeated-measures ANOVA with type of combination (human–animal, human–plant, human–object, animal–plant, animal–object, plant–object) as the independent variable and RT as the dependent variable. Mean RT for correct responses and accuracy rate are shown in Table 7. The RT analysis showed no significant differences among the six combinations ($F(5,125) = 1.60$, n.s.). In addition, the percent of correct responses did not differ across the four categories ($F(5,130) = 1.17$, n.s.).

4.3. MIXED ANALYSIS

We performed a mixed-model analysis with subjects and items as random effects using SPSS 15.0 software for Windows (SPSS Inc, Chicago, IL). The dependent variable (RT) was not normally distributed therefore analysis was performed on log (RT) of correct responses. Compound symmetry was assumed between items. Responses that were slower than two standard deviations from the mean RT of each participant were removed from the analysis.

There was a significant condition effect ($F(4,2090) = 12.40$, $p < .0001$). Results did not reveal a hierarchy effect, indicating that there was no significant difference in RT between the congruent hierarchically-compatible (923 ms) and the congruent hierarchically-incompatible (951 ms) conditions ($t(2089) = -0.60$, n.s.). Additional results indicated that RT in the congruent hierarchically-incompatible (951 ms) were shorter than were RT in the incongruent hybrid condition (1002 ms) ($t(2089) = 4.38$, $p < .001$). Also, participants responded more quickly in the congruent hierarchically-incompatible hybrid condition (951 ms) than in the congruent neutral condition (1030 ms) ($t(2089) = 4.73$, $p < .001$).Thus, in contrast to the results of

Experiment 1 that showed a hierarchy effect when hybrids were matched to grammatically asymmetrical phrases, no such effect was found when grammatically symmetrical constructions were used.

4.4. COMPARING THE HIERARCHY EFFECT OF EXPERIMENTS 1 AND 3

To calculate the hierarchy effect, we subtracted every person's RT on the congruent hierarchically-compatible true hybrid condition from that person's RT on the congruent hierarchically-incompatible true hybrid condition and then calculated the mean difference across participants. We used an independent-samples t-test to compare the hierarchy effect in Experiment 1 (mean = 87 ms, $SD = 107.34$) and in Experiment 3 (mean = 28, $SD = 74.50$). This test showed a significantly greater condition effect for grammatically asymmetrical phrases relative to grammatically symmetrical constructions ($t(47) = 2.31$, $p < .05$).

5. General discussion

Two major findings emerged from the present set of studies. First, there was a difference between the verbal task (Experiment 1) and the visual task (Experiment 2), so that a hierarchy effect was documented only in the language-mediated task. Second, there was a difference between grammatically asymmetrical phrases (Experiment 1) and grammatically symmetrical phrases (Experiment 3), so that the hierarchy effect was observed only in grammatically asymmetrical constructions. Furthermore, the hierarchy effect was greater in the grammatically asymmetrical condition than in the corresponding symmetrical condition (Experiment 1 vs. Experiment 3).

The difference between Experiments 1 and 2 suggests that the conceptual hierarchy does not apply to non-verbal categorization, either because visual hybrids are not conceived of as entities in the first place (see Wagner et al., 2010) or because they constitute totally new entities. However, when asked to connect the conceptualization of this new entity to a verbal description, particularly to grammatically asymmetrical constructions, people rely on the same principles that guide their processing of more familiar stimuli. We argue that hierarchically compatible phrases trigger the pragmatic tendency to infer that nouns that have a higher grammatical function (e.g., the subject or the head noun) should also be higher on the conceptual hierarchy (Feleki, 1996; Branigan et al., 2008; Shen & Gil, 2010). Therefore, under time constrains, matching a hybrid to a grammatically asymmetrical phrase is facilitated by hierarchically compatible phrases relative to hierarchically incompatible phrases.

Another potential explanation for the order effects that were found in the linguistic task in Experiment 1 may relate to collocation frequency. If so, than perhaps the fact that hybrids tend to be coded this way in language is indicative of ‘deeper’ underpinnings, which gave rise to the preferred collocation order. However, when Branigan et al. (2008) tested the CH effect in symmetrical and asymmetrical constructions, they found no effect in the symmetrical constructions. Consistent with these findings, our results suggest that the conceptual hierarchy effect emerges only in grammatically asymmetrical phrases. If the hierarchical effect relies on a rudimentary underlying mechanism then it should be present in both the grammatically symmetrical and the grammatically asymmetrical phrases because the same pairs of nouns (e.g., *man*, *lizard*) appear in the two types of phrase. Yet our findings suggest that the effect appears only in asymmetrical phrases such as *a man with a lizard tail* but not in symmetrical ones such as *half-man half-lizard*.

Note that both verbal and pictorial uses of metaphors are also sensitive to the conceptual hierarchy. For example, Connor and Kogan (1980) as well as Kogan et al. (1989) showed that when participants describe picture triads in which two pictures are metaphorically related (e.g., a picture of a rosebud, a baby, and a watering can), they use an asymmetrical metaphorical comparison (A *is like* B). Thus, pictures that represent a hierarchically higher entity (e.g., the baby) are more likely to serve as the topic of the comparison, and pictures that represent a hierarchically lower entity (e.g., the rosebud) serve as the source. A similar pattern was found when the stimuli consisted of verbal representations of these visual figures. These findings have been interpreted as evidence in support for a pre-linguistic (conceptual) hierarchy, since the CH effect was observed when both linguistic stimuli (word triads) and non-linguistic stimuli (picture triads) were used. However, since both tasks involved the use of language, it is unclear whether these results actually demonstrate that pictures are conceptually categorized in terms of the CH.

When a hybrid does not match the preceding verbal phrase, as in our incongruent conditions, there is no difference between responses to hybrids preceded by hierarchically-compatible phrases and responses to hybrids preceded by hierarchically-incompatible phrases. That is, when the pragmatic mapping is irrelevant because no matching can be performed, participants do not show a CH facilitation effect. Thus, hybrids are not ‘transparent’ and differences in processing times may be attributed to the pragmatic mapping and not to the type of phrase. The analysis of error rate for matching hybrids with verbal descriptions in Experiment 1 provides further support for this account. Although accuracy level was very high (Table 2), as the task was simple, error rate was almost three times greater in conditions that included hierarchically-incompatible phrases (6.15%) than in conditions that included hierarchically-compatible phrases (2.38%). It appears that under time constraints, participants are more sensitive to the nature of the verbal phrase than to the visual image. This might also be the reason why we found a difference between the two types of verbal phrase used in Experiments 1 and 3. While grammatically asymmetrical constructions yielded a hierarchy effect, no such effect was found in the grammatically symmetrical condition.

Thus, according to the present account, the difference between the two types of verbal construction used in Experiments 1 and 3 shows that the hierarchy effect reflects a linguistic tendency that applies to the asymmetrical constructions alone. This interpretation is in line with the lack of CH effect in ordering nouns within noun phrase conjunctions. Indeed, both Tanaka et al. (2005) and Branigan et al. (2008) reported that the CH does not affect the ordering of nouns (or noun phrases) in noun-phrase conjunctions. The use of a verbal task, then, led to a hierarchy effect in the asymmetrical construction, but not in the symmetrical construction.

This pattern is consistent with the notion of 'thinking for speaking' (Slobin, 1996), according to which thoughts are converted into lexical categories and grammatical structures, thus changing preverbal conceptualization as we use language (e.g., in speaking, writing, or reading). According to Slobin: "the activity of thinking takes on a particular quality when it is employed in the activity of speaking" (1996, p. 76); that is, 'thinking' adopts different forms depending on the mode that we use (verbal or visual mode).

Note that 'thinking for speaking' differs from 'habitual thinking', the stronger Whorfian hypothesis, according to which one's language determines the long-lasting habitual conceptualization of reality, rather than the short-lived effects that emerge as we use language. Whereas Whorf argues that thinking is affected by both linguistic and conceptual tasks (Whorf & Chase, 1956), the 'thinking for speaking' hypothesis maintains that language does not have a long-term influence on thinking, but the effect is much more restricted. The assumption that the effect is temporary and occurs as we speak is fully compatible with our findings that the hierarchy effect was not present on the non-verbal task. Had the linguistic tendency influenced habitual thought, the CH effect would occur regardless of whether the categorized stimuli were verbal or non-verbal. Instead, only on-line use of language resulted in the CE effect.

The 'thinking for speaking' theory has been developed within the 'linguistic relativity' framework (see Wolff & Holmes, 2011, for a recent review) in order to explain why speakers of diverse languages think differently. Note, however, that the notion of 'thinking for speaking' presupposes a difference between 'thinking' and 'thinking for speaking' for the SAME linguistic community. As Wolff and Holmes (2011) put it: "Most work investigating the relationship between language, cognition, and perception has assumed that verbal and non-verbal representations are fundamentally distinct and the goal of the 'language and thought' research program is to understand whether and how linguistic representations affect non-linguistic representations."

In sum, the present study suggests that the conceptual principles that guide verbal processes do not necessarily guide the processing of visual hybrids, so that the effect of these principles is limited to linguistic stimuli that involve grammatical asymmetry.

REFERENCES

Abboud, H. (1991). SuperLab. Wheaton, MD: Cedrus.

Boroditsky, L., Schmidt, L., & Phillips, W. (2003). Sex, syntax, and semantics. In D. Gentner & S. Goldin-Meadow (Eds.), *Language in mind: advances in the study of language and thought* (pp. 61–80). Cambridge, MA: MIT Press.

Branigan, P. H., Pickering, M. J., & Tanaka, M. (2008). Contributions of animacy to grammatical function assignment and word order during production. *Lingua*, **118** (2), 172–189.

Carroll, N. (1994). Visual metaphor. In Jaakko Hintikka (Ed.), *Aspects of metaphor* (pp. 189–218). Dordrecht: Kluwer.

Connor, K., & Kogan, N. (1980). Topic–vehicle relations in metaphor: the issue of asymmetry. In Richard P. Honeck, & Robert. R. Hoffman (Eds.), *Cognition and figurative language* (pp. 283–310). Hillsdale, NJ: Erlbaum.

Cooper, W. E., & Ross, J.R. (1975). Word order. In R. G. Grossman, L. San, & T. Vance (Eds.), *Papers from the parasession on functionalism* (pp. 63–111). Chicago: Chicago Linguistic Society.

de Villiers, J. G., & de Villiers, P. A. (2003). Language for thought: coming to understand false beliefs. In D. Gentner, & S. Goldin-Meadow (Eds.), *Language in mind: advances in the study of language and thought* (pp. 335–386). Cambridge, MA: MIT Press.

Deane, D. P. (1992). *Grammar in mind and brain: explorations in cognitive syntax*. Berlin: Mouton de Gruyter.

Deane, D. P. (1993). On metaphoric inversion. *Metaphor and Symbolic Activity*, **8** (2), 111–126.

Feleki, E. (1996). *The effects of conceptual accessibility on language production: experimental evidence from Modern Greek*. (Unpublished MA thesis) Centre for Cognitive Science, University of Edinburgh.

Forceville, C. (1996). *Pictorial metaphor in advertising*. London & New York: Routledge.

Gentner, D., & Goldin-Meadow, S. (Eds.) (2003). *Language in mind: advances in the study of language and thought*. Cambridge, MA: MIT Press.

Gleitman, L., & Papafragou, A. (2005). Language and thought. In K. J. Holyoak, & R. G. Morrison (Eds.), *The Cambridge handbook of thinking and reasoning* (pp. 633–661). Cambridge: Cambridge University Press.

Keil, F. C. (1979). *Semantic and conceptual development: an ontological perspective*. Cambridge, MA: Harvard University Press.

Keil, F. C. (1981). Constraints on knowledge and cognitive development. *Psychological Review*, **88** (3), 197–227.

Kogan, N., Mindi, C., & Heleen, H. (1989). Developmental trends in metaphoric asymmetry. *Metaphor and Symbolic Activity*, **4** (2), 71–91.

Levin, B., & Hovav, M. R. (2005). *Argument realization*. New York: Cambridge University Press.

Loewenstein, J., & Gentner, D. (2005). Relational language and the development of relational mapping. *Cognitive Psychology*, **50**, 315–353.

Lupyan, G. (2012). Linguistically modulated perception and cognition: the label-feedback hypothesis. *Frontiers in Psychology*, **3** (54), 1–13.

Lupyan, G., Rakison, D. H., & McClelland, J. L. (2007). Language is not just for talking: redundant labels facilitate learning of novel categories. *Psychological Science*, **18**, 1077–1083.

Merchant, J. (2006). Polyvalent case, geometric hierarchies, and split ergativity. *Proceedings from the Annual Meeting of the Chicago Linguistic Society* (**42** (2), 57–76). Chicago: Chicago Linguistic Society.

Quene, H., & van den Bergh, H. (2008). Examples of mixed-effects modeling with crossed random effects and with binomial data. *Journal of Memory and Language*, **59**, 413–25.

Roberson, D., Davidoff, J., Davies, I. R., & Shapiro, L. R. (2005). Color categories: evidence for the cultural relativity hypothesis. *Cognitive Psychology*, **50**, 378–411.

Shen, Y., & Gil, D. (2010). *The perception of visual hybrids: the role of language*. Paper presented at Cognitive Poetics and Rhetoric 1.0, University of Lodz, Poland.

Shen, Y., Gil, D., & Roman, H. (2006). What can hybrids tell us about the relationship of language and thought? Paper presented at the IGEL conference. Munich, Germany.

Slobin, D. (1996). From 'thought and language' to 'thinking for speaking'. In John J. Gumperz, & Stephen C. Levinson (Eds.). *Rethinking linguistic relativity* (pp. 70–96). Cambridge: Cambridge University Press.

Tanaka, M., Holly, P. B., & Martin, J. P. (2005). The role of animacy in Japanese sentence production. Paper presented at the CUNY Sentence Processing Conference, Tucson, AZ.

Wagner, W., Kronberger, N., Nagata, M., Sen, R., Holtz, P., & Palacios, F. F. (2010). Essentialist theory of 'hybrids': from animal kinds to ethnic categories and race. *Asian Journal of Social Psychology*, **13** (4), 232–246.

Waxman, S., & Markow, D. B. (1995). Words as invitations to form categories: evidence from 12- to 13-month-old infants. *Cognitive Psychology*, **29**, 257–302.

Whorf, B. L., & Chase, S. (1956). *Language, Thought and Reality, Selected Writings of Benjamin Lee Whorf. Edited… by John B. Carroll. Foreword by Stuart Chase*. J. B. Carroll (Ed.). Mass.

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