When Skunks are Similar to Giraffes and when they are not: Grammatical Gender Effects on Bilingual Cognition

Stavroula-Thaleia Kousta (s.kousta@ucl.ac.uk),
Department of Psychology, 26 Bedford Way
London WC1H 0AP United Kingdom

David P. Vinson (d.vinson@ucl.ac.uk),
Department of Psychology, 26 Bedford Way
London WC1H 0AP United Kingdom

Gabriella Vigliocco (g.vigliocco@ucl.ac.uk)
Department of Psychology, 26 Bedford Way
London WC1H 0AP United Kingdom

Abstract

We investigate how second language learning affects 'thinking for speaking' in the first and second language by examining the semantic effects of grammatical gender (present in Italian but absent in English). In an error-induction experiment we first establish a baseline, showing that gender affects the semantic substitution errors made by monolingual Italian speakers compared to monolingual English speakers. We then show that Italian—English bilinguals behave like monolingual English speakers when the task is in English, and like monolingual Italian speakers when the task is in Italian, hence exhibiting appropriate 'thinking for speaking' for each language. These results have implications for linguistic relativity/determinism and models of bilingual semantic memory and processing.

Keywords: bilingualism; linguistic relativity; thinking for speaking; grammatical gender; speech errors

Introduction

In order to speak any language, one has to pay attention to the distinctions obligatorily expressed in the language, and in this way every linguistic community differs from every other. In other words, there is linguistic relativity, since speakers of typologically different languages are required to verbalize different aspects of reality when constructing linguistic messages. What are the cognitive consequences of linguistic relativity? For Whorf (1956), to whom the linguistic relativity hypothesis is attributed, linguistic relativity is closely linked to linguistic determinism, that is, the proposal that language determines the way the external world is perceived, categorized and acted upon.

Nowadays virtually nobody would like to claim that language has such a deterministic role on cognition as Whorf originally envisioned it. It is, however, still a matter of intense debate to what extent language affects (rather than determines) cognition. According to one hypothesis, the linguistic classifications imposed by language affect only those aspects of cognition that are actively engaged in the processes of speaking. This "thinking for speaking" proposal has been put forward by Slobin (1996), is currently more or less uncontroversial, and evidence in support of it abounds in the literature (see Vigliocco, Vinson, Paganelli, & Dworzynski, 2005).

A stronger version of the linguistic determinism hypothesis argues that language also affects non-linguistic cognition, above and beyond the processes involved in speaking and comprehending a particular language (e.g., speakers of different languages perceive the world in different ways due to linguistic differences). Data in support of this claim have been provided in numerous behavioral experiments (e.g. Imai & Gentner, 1997; Levinson, 1997; Lucy, 1992), but at the same time there is a substantial body of evidence that argues against this claim (e.g. Gennari et al., 2002; Li & Gleitman, 2002; Vigliocco et al., 2005). In other words, the main debate in the field focuses on the extent and pervasiveness of the effect of language on cognition.

Up to the present, investigations of the link between language and thought have concentrated on monolingual populations, with very little work carried out with bilingual speakers. Bilingualism, however, offers a unique test case for the potential role of language in shaping cognition, especially when focusing on highly proficient bilingual speakers who acquired a second language after their first language was largely in place. If learning a first language affects cognition beyond 'thinking for speaking' for that same language, then bilingual speakers, even when they are highly fluent in their second language, should show evidence of transfer of linguistically motivated semantic/conceptual categories from their first language into their second, despite their high proficiency in their second language. In other words, their 'thinking for speaking' in their second language should be, at least to some extent, dependent on 'thinking for speaking' in the first. This prediction is based on the assumption that language, by means of its lexical and grammatical structure, habitually directs attention to a number of conceptual distinctions.

In this paper we investigate bilingual cognition in relation to a phenomenon, grammatical gender, which has attracted considerable interest in the literature on linguistic relativity/determinism. The crucial questions are whether the bi-

lingual speakers' behavior in each of their languages converges or diverges, how it relates to the behavior of monolingual speakers, and whether any effects in linguistic tasks extend to non-linguistic tasks. These questions have not yet been satisfactorily addressed in the literature, but they are extremely important in order to understand the extent and nature of the influence of language on cognition.

In many languages, grammatical gender is a formal category: nouns are marked masculine, feminine, neuter etc.. Italian, for instance, has two genders, and all nouns are either feminine or masculine; while in English gender is not a formal category. Does the grammatical gender of words affect their semantic representation and the conceptual representation of their referents? Is the meaning of *giraffe* more similar to the meaning of *skunk* for Italian speakers than English speakers just because the words share grammatical gender in Italian (giraffa$_{fem}$—puzzola$_{fem}$)? And are zebras in Italian perceived as being feminine and leopards perceived as masculine just by virtue of the grammatical gender of the words used to refer to them (giraffa$_{fem}$—leopardo$_{masc}$)?

Early work with monolinguals suggested that grammatical gender affected both the semantic representation of words and the conceptual representation of their referents (see Sera et al., 2002 for discussion). More recent research, however, has suggested that grammatical gender effects are substantially constrained. Sera, et al. (2002) asked speakers to assign male or female voices to pictured objects, and observed gender effects with speakers of languages such as French and Spanish, but not German, despite the fact that all three languages have formal gender. The reason for this was proposed to be the fact that Spanish and French have a two-gender system (masculine-feminine) which exhibits a high correlation between grammatical and natural gender, and gender is morphologically marked across several grammatical categories. German, on the other hand, has a three-gender system (masculine-feminine-neuter) with a less straightforward relationship between grammatical and natural gender, and gender is marked in fewer grammatical categories. It is possible, however, that gender effects were observed in Spanish and French where speakers could successfully utilize gender as a strategy to assign male and female voices to pictures, while with German speakers the strategy fails with nouns that are marked neuter, hence leading to a null result.

Using tasks less susceptible to the strategic use of gender, Vigliocco et al. (2005) replicated the difference between two- and three-gender languages and showed that gender effects are even further constrained: apart from being limited to languages with two genders and with a close correspondence between grammatical and natural gender, grammatical gender was shown to affect the semantic representation of words referring to sexuated entities (e.g. animals) but not the representation of words referring to entities that lack natural gender (e.g. artifacts). Moreover, this constrained semantic effect did not generalize to a non-linguistic task (triadic judgments with pictures).

In the present study we explore a domain in which lan-

guage-specific semantic effects of grammatical gender have been clearly observed (Italian words referring to animals), in order to investigate to what extent bilingual speakers adapt to 'thinking for speaking' that is appropriate for their second language; to what extent 'thinking for speaking' in a second language is affected by the first; and to what extent 'thinking for speaking' in the first language is affected by learning a second language. To address these questions, we focus on semantic errors arising during continuous naming, an online methodology that is not subject to the use of strategies related to the effects of gender (see Vigliocco et al., 2005). We contrast the errors that are produced in Italian to those produced in English, a language without grammatical gender. Comparing a gendered to a non-gendered language allows us a) to focus specifically upon whether grammatical gender actually affects semantic representations in processing (since the monolingual English data provide a baseline of semantic/conceptual/visual similarity that is shared among speakers of the two languages) and b) to determine whether this semantic effect is transferred to a language that lacks the specific grammatical category.

We ask monolingual Italian and monolingual English speakers to name pictures of common land animals that are presented at a fast rate. Bilingual Italian-English speakers are also asked to carry out the same task in both their languages. The aim of the task is to elicit semantic substitution errors (e.g. saying 'eye' when 'ear' is intended), which are assumed to arise as a result of competition between semantically related lexical candidates in a conceptually driven lexical selection process during production (Garrett, 1984). In the literature on language production, it is commonly assumed that grammatical gender, a lexico-syntactic property, affects the production of phrases, but not the production of single words (Schriefers et al., 2002). However, for languages like Italian and semantic fields like animals (which have natural gender), Vigliocco et al. (2005) have shown that gender is also a lexico-semantic property, affecting the production of single words. In order to investigate whether grammatical gender effects are only due to the activation of syntactic information necessary for computing agreement, we compare errors in naming single words to errors in producing simple noun phrases (determiner+noun; Italian determiners are marked for gender).

If grammatical gender increases the semantic similarity between nouns that share gender, then the errors produced by Italian speakers should reflect not only the similarity in meaning that is evident in the errors produced by monolingual English speakers, but also additional effects of grammatical gender (as found by Vigliocco et al., 2005). In Experiment 1 we replicate the results of Vigliocco et al. (2005), and conduct additional analyses in order to establish whether gender effects are still observed when multiple, fine-grained measures of phonological similarity are also taken into account. In Experiment 2 we investigate the behavior of bilingual speakers to see if grammatical gender in their first language affects performance in their second language; whether the effect is restricted to their first language

only; or whether learning a second language without formal gender has consequences for first-language processing.

Experiment 1: Monolingual error induction

Method

Participants Twenty-six native English speakers were recruited through the UCL psychology subject pool, and were paid at a rate of £6/hour. Twenty-five native Italian speakers (psychology students at the Università degli Studi di Trieste, Italy) also participated and were paid at a rate of €9/hour or received course credit. No English participants reported moderate or higher competence in any gendered language; no Italian participants reported moderate or higher competence in any second language.

Materials and design The set of items consisted of pictures depicting common land animals. They were the same items used for Experiment 2 in Vigliocco et al. (2005): pictures of 27 animals (11 marked feminine in Italian; 16 masculine).

Forty-one blocks of ten pictures were created by randomly selecting pictures, observing the following constraints: first, a picture appeared no more than once within a block, and second, a picture never appeared as the last item in one block and the first item in the next. Each picture appeared either 15 or 16 times in each condition and each participant saw a different random order of blocks for each part of a session.

Procedure Participants were told that they were taking part in a study of speech patterns under time pressure, and that they would be asked to name pictures depicting animals as they appeared on screen. The experiment was carried out in two parts: in one part participants were asked to name the pictures using only a single word, while in the other part they were asked to produce noun phrases consisting of a definite article and noun.

Speakers were asked to attempt to keep up with the rate of presentation, i.e. naming pictures as they appeared, rather than retaining them in memory, skipping items if necessary to recover from difficulty. All spoken responses were digitally recorded and later transcribed and scored.

The experiment began with an untimed name agreement phase, in which each picture was presented, and participants were asked to name them without time pressure. After this, participants performed a set of practice blocks. These practice blocks were intended not only to familiarize participants with the task, but also to allow the experimenter to adjust the rate of presentation according to each participant's performance. After each block of 10 pictures, the experimenter altered the rate of presentation in order to make the task difficult but manageable for each speaker. Final presentation rates ranged between 400 and 900 ms display time for English (average=633ms) and between 500 and 1200 ms for Italian (average=688ms).

Once the practice session was completed, the experimental blocks were presented. Participants pressed a key to begin each block, and then the 10 items in the block appeared in sequence at randomly-selected positions on the screen, with time parameters as determined in the practice session.

Results

The data were transcribed and each response was scored as belonging to one of the following categories: *correct*: the intended noun was produced in its entirety (85.0% of all responses); *different label*: participant consistently produced a different noun than intended by the experimenters, but which could be considered an acceptable response (4.3%); *lexical error*: a label was produced that did not qualify as correct or a different label (2.7%); *omission*: no response was made for a given picture (6.5%); *self-correction*: participant started producing an incorrect word, but changed their response to a different word before it was complete (0.4%). These were treated as lexical errors if the incorrect word was produced completely before the correction was made; *miscellanea*: dysfluencies, incomplete/inaudible responses, etc. (1.1%). Only responses classified as lexical errors were included in the analyses reported below.

The likelihood of substituting one word with another is increased if target and intruder share phonological, in addition to semantic, similarity (the mixed error effect; Dell & Reich, 1981). It is important to exclude such errors from the data set, because there are clear phonological correlates of grammatical gender in Italian. In order to eliminate the confounding effect of phonological similarity, we carried out the following procedure. First, we phonologically transcribed all the animal nouns using standard British English and standard Italian pronunciations. We then excluded all target-intruder pairs that shared more than 33% of their phonemes in the same position (onset, nucleus, coda of the first syllable, middle syllable(s), and the final syllable). This was also done individually for participants with accents that differed from the standard. After applying these procedures, there remained 393 English errors and 344 Italian errors. These errors can be considered semantic, rather than mixed.

For each participant and item we calculated the proportion of these errors for which the target and the intruder shared Italian gender.[1] Table 1 presents the average proportion of target-error pairs sharing gender in the two languages and response types.

Table 1: Mean proportion of gender-preserving errors by monolingual English and Italian speakers in single word and noun phrase naming conditions (95% CI in brackets).

Language	Single word	Noun phrase
English	.32 (.21-.44)	.38 (.26-.49)
Italian	.66 (.55-.78)	.61 (.49-.73)

[1] English words were assigned gender based on their translation-equivalent words in Italian. Because English does not have grammatical gender, the English "gender preservation rate" indicates the rate of gender preservation that would be expected in Italian if gender had no language-specific effects on the resulting errors.

We performed 2 x 2 ANOVAs with participants (F_1) and items (F_2) as random factors, investigating the effects of language (English, Italian—between subjects, within items) and response type (single word, noun phrase—within subjects and items). There was a significant main effect of language in both analyses ($F_1(1, 45)=19.009$, $p<.001$, $MS_e=.065$; $F_2(1,19)=13.300$, $p<.01$, $MS_e=.057$), with Italian errors sharing gender more often than English errors ($M_{difference}=.28$, 95% CI=.13). There was no main effect of response type or interaction. These results show that grammatical gender affects the perceived semantic similarity between different animals for Italian speakers as compared to English speakers (replicating Vigliocco et al., 2005).

However, it is possible that the selection of an arbitrary criterion of phonological similarity, based only on shared segments all treated equally, may not rule out all aspects of phonological similarity that are correlated with shared gender among the items tested. Therefore we conducted an additional analysis, using multiple regression to consider whether gender actually affects the errors produced by Italian speakers once a number of potential confounding phonological factors are taken into account. Because the vast majority of lexical errors were also from within the response set, we began by creating a confusion matrix by pairing all targets with all possible errors, assigning a value of 1 for target-error pairs that were observed in the dataset (120 distinct target-error pairings), and 0 otherwise. The Italian errors treated in this manner served as the dependent variable in this analysis.

The regression included three steps. In the first step, the presence or absence of errors in the English dataset (121 observed target-error pairings) was used to predict presence or absence of errors in Italian. This takes into account those aspects of similarity among animals that are common between English and Italian, such as conceptual similarity and visual similarity among the pictures used in the experiment. The residuals from this model were then passed to a second step: a variety of predictors related to phonological similarity in Italian. These included whether the two words shared the same number of syllables, stress pattern, onset phoneme, final phoneme, and also the continuous measure of form similarity described above (% of phonemes shared in the same position between target and error). Finally, the residuals of this second step were passed to a final step, in which only shared grammatical gender was used as a predictor. Taking such an approach, common variance due to correlations between gender and morphophonological similarity, semantic similarity or visual similarity has been separated out in the initial steps. Therefore, if gender remains a significant predictor in the final step, it can be said that gender *per se* is predicting the presence of errors.

Unsurprisingly, the presence or absence of errors in English was a strong predictor of Italian errors; $r_{partial} = .336$, $t = 9.409$, $p < .001$. In the second step, only the fine measure of phonological similarity (% segments shared by position) was a significant predictor; $r_{partial} = .107$, $t = 2.843$, $p = .005$ (all other form predictors $p > .3$). Crucially, in the final step gender remained a significant predictor of errors; $r_{partial} = .137$, $t = 3.645$, $p < .001$ (final model $R^2 = .157$). This shows that the gender effect observed in the ANOVA is not simply a product of the arbitrary cutoff at which errors were decided to be "phonologically similar".

Discussion

Experiment 1 demonstrates that grammatical gender affects perceived semantic similarity in Italian as compared to English (a language with no formal gender), replicating the results in Vigliocco et al. (2005). The monolingual Italian speakers' pattern of errors differed from that of the monolingual English speakers: a significantly higher proportion of gender-preserving errors even after semantic and phonological correlates of gender were factored out. Grammatical gender, essentially a lexico-syntactic property, appears to have acquired semantic status in Italian (at least for words referring to animals); words' meanings are affected by their grammatical specification (masculine vs. feminine).

One result still requires explanation: the proportion of gender-preserving errors in Italian was the same for single words and noun phrases, in contrast to previous work (e.g. Paganelli et al., 2003) where higher gender preservation rates in Italian have been reported for noun phrase production. According to one account (e.g., Costa & Caramazza, 2002), grammatical gender effects in noun phrase production are due to determiner selection which occurs during the phonological encoding of the noun phrase in languages like Italian. Because determiner selection occurs so late it has little effect on lemma selection. Alternatively, gender preservation effects would arise as a result of monitoring at the morphophonological level: lemmas competing for selection also activate syntactic frames specified for gender (see Vigliocco et al., 2004). When competition is high, a frame might be selected before a lemma is selected. Gender mismatch between frame and lemma would then be more easily detected by a monitoring system. In the present experiment, the high level of competition among same-gender lemmas (being more semantically similar because of shared gender) may cause fewer ill-formed syntactic frames to be generated in the first place. At any rate, resolving this issue is beyond the scope of the present paper and it does not affect the main purpose of this data set, which is to serve as a baseline for bilingual speakers.

This leads to the main question addressed in this paper: the extent to which bilingual speakers adapt to thinking for speaking appropriate for their second language and the extent to which the representations of their mother tongue are affected by learning a second language. This is addressed in Experiment 2, in which bilingual speakers are asked to name pictures in Italian and in English.

Experiment 2: Bilingual error induction

Method

Participants Twenty-nine native Italian speakers, all highly proficient in English, were recruited for the experiment. All had learnt English after the age of 6, had an average of 7.4 years of formal instruction in English, and all had been resident in the UK for at least 9 months prior to the time of testing (mean length of stay in the UK = 4.6 years). Volunteers were paid at a rate of £12/hour for their participation.

Materials, design and procedure These were the same as those in Experiment 1, with the following exceptions. Bilingual speakers did the same experiment in English and Italian (on different days). The order of the two language sessions was counterbalanced across participants to avoid any order effects. The order of the single word and the noun phrase conditions was also counterbalanced across participants.

Results

The data were recorded and transcribed as before. *Correct* responses made up 86.6% of all responses; *different label*, 3.9%; *lexical error*, 2.5%; *omission*, 7.0%; *self-correction*, 0.6% and *miscellanea*, 1.2%. Once again, in order to make possible the comparison between English and Italian, English words were assigned the gender that translation equivalent items have in Italian. The proportion of gender-preserving lexical errors as a function of language and response type appears in Table 2.

Table 2: Mean proportion of gender-preserving errors by bilingual speakers in English and Italian, in single word and phrase naming conditions (95% CI in brackets).

Language	Single word	Noun phrase
English	.35 (.22-.49)	.35 (.22-.47)
Italian	.59 (.47-.71)	.62 (.48-.77)

We analyzed all lexical errors in the same manner as in Experiment 1, including the exclusion of mixed errors. We also excluded errors from both languages in which the target and the *translation* of the error shared 50% or more of their phonemes in the same serial order, disregarding syllable boundaries and structure.

After mixed errors had been excluded, there remained 347 Italian errors and 283 English errors, which we analyzed in the same manner as in Experiment 1. Only participants who contributed observations to both conditions in both languages were included in the analyses. The analyses yielded the same pattern of results as for the monolingual speakers: a significant main effect of language ($F_1(1,23)$=13.292, p<.001, MS_e=.070; $F_2(1,20)$=7.499, p<.05, MS_e=.042), with more gender-preserving errors in Italian than in English ($M_{\text{difference}}$=.25, 95% CI=.14). There was no main effect of response type or interaction (all Fs<1). Bilingual speakers in this experiment were significantly more likely to produce same gender intruders when they were carrying out the task in Italian than the same task in English.

As in Experiment 1, we explored this grammatical gender effect with regression analyses. We tested the Italian bilingual errors using the English bilingual errors as a predictor in the first step. Italian form similarity was also included as before, and gender in a final step. Errors in English were a strong predictor in the first step (r_{partial} = .388, t = 11.106, p < .001). In the second step, fine-grained phonological similarity was the only form-based predictor to reach significance (r_{partial} = .108 = 2.857, p = .004). Gender was also a significant predictor in the final step (r_{partial} = .399, t = 3.778, p < .001; combined model R^2 = .198), showing that bilingual speakers' errors preserved grammatical gender when they were responding in Italian.

Finally, to determine the extent to which the performance of the bilingual speakers was equivalent to that of the monolingual speakers, we carried out two further sets of ANOVAs assessing the effects of language (bilingual and monolingual—between participants but within items) and response type (single word and noun phrase—within participants and items) for each of the two languages. None of the main effects or interactions reached significance in these analyses (also largely confirmed by regression analyses), indicating that bilingual speakers' semantic representations for these items do not differ from those of monolingual speakers, depending on the language in which the task is carried out.

Discussion

Experiment 2 showed that bilingual speakers made more same-gender errors in Italian than for the same items in English. Moreover, we did not find any significant difference between the bilingual and the monolingual populations. We interpret these results as evidence for the fact that the bilingual speakers in these studies demonstrate 'thinking for speaking' behavior that is appropriate for each of their languages. In other words, there is intra-speaker linguistic relativity and the behavior of bilingual speakers in a given language is predicted by the behavior of monolingual speakers.

Evidence for linguistic relativity based on comparisons of monolingual speakers has clearly different consequences for claims of linguistic determinism than evidence for linguistic relativity with bilingual speakers. In the monolingual literature, differences between linguistic groups on a specific semantic/conceptual variable have been interpreted as indicating the potentially central role language has on shaping cognition. Evidence for linguistic relativity within the same speaker (the bilingual individual) can only have the reverse interpretation: if proficient bilinguals who acquired their second language after their first show evidence of adaptation to monolingual norms for their second language, their first language has a very limited effect, an effect on 'thinking for speaking' just for that language.

General Discussion

Our aim has been to determine to what extent proficient bilingual speakers adapt to 'thinking for speaking' that is appropriate for their second language or not and to what extent learning a second language affects 'thinking for speaking' in the first. The results obtained show that bilingual L1 Italian-L2 English speakers' performance was significantly different in each of their languages, thus demonstrating relativity in their 'thinking for speaking'. Moreover, we did not find any evidence of transfer from L1: there was no significant difference between the monolingual and the bilingual English data. Finally, learning a second language without grammatical gender does not seem to affect semantic representations in the gendered first language, since there was no significant difference between the monolingual Italian and bilingual Italian data.

Evidence of transfer from one's native/most proficient language into a second language has been interpreted as evidence for the pervasive effect of language on cognition, especially when highly proficient bilinguals are involved (Boroditsky, Schmidt, & Phillips, 2003). We found no evidence of transfer from Italian into English of the semantic effects of gender and interpret this lack of transfer as evidence for the constrained role grammatical gender has on bilingual cognition. If bilingual speakers are able to adapt to 'thinking for speaking' that is appropriate for a second language, even when they are not as proficient in their second language as in their first, this constitutes evidence for a limited cognitive effect of the particular variable both in the first and the second language.

Bilingualism and multilingualism are the norm rather than the exception around the world (Harris & McGhee Nelson, 1992). Despite this, research on language-specific effects on cognition has focused on monolingual populations, not addressing how two or more different ways of talking about the world are accommodated within a single mind/brain. Such research is essential not only in order to understand the cognitive consequences of bilingualism but also in order to determine the extent of the effects of language on cognition. One test of the pervasiveness of language-specific effects on cognition is its effect on the representations of a second language. The present paper addressed this question and concluded negatively, at least for the specific phenomenon studied. These results also make it highly unlikely that Italian grammatical gender has an effect on the non-linguistic, conceptual representations of bilingual speakers. Future investigation will need to address a wider variety of linguistic phenomena in order to determine the nature of the link between linguistic variability and bilingual cognition.

Acknowledgments

References

Boroditsky, L. Schmidt, L.A., & 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*. Cambridge, MA: MIT Press.

Costa, A. & Caramazza, A. (2002). The production of noun phrases in English and Spanish: Implications for the scope of phonological encoding during speech production. *Journal of Memory and Language, 46*, 153-177.

Dell, G.S., & Reich, P.A. (1981). Stages in sentence production: An analysis of speech error data. *Journal of Verbal Learning and Verbal Behavior*, *20*, 611-629.

Garrett, M. F. (1984). The organization of processing structure for language production: Application to aphasic speech. In D. Caplan, A. R. Lecours and A. Smith (Eds.), *Biological perspectives on language*. Cambridge: MA: MIT Press.

Gennari, S.P., Sloman, S., Malt, B., & Fitch, T. (2002). Motion events in language and cognition. *Cognition, 83*, 49-79.

Harris, R.J. & McGhee Nelson, E.M. (1992). Bilingualism: Not the exception any more. In R.J. Harris (Ed.), *Cognitive processing in bilinguals*. Amsterdam: North-Holland.

Imai, M. & Gentner, D. (1997). A crosslinguistic study of early word meaning: Universal ontology and linguistic influence. *Cognition, 62*, 169-200.

Levinson, S. C. (1997). From outer to inner space: linguistic categories and non-linguistic thinking. In J. Nuyts & E.Pederson (Eds.), *Language and Conceptualization*. Cambridge: Cambridge University Press

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

Lucy, J.A. (1992). *Grammatical categories and cognition: A case study of the Linguistic Relativity Hypothesis*. Cambridge, UK: Cambridge University Press.

Paganelli, F., Vigliocco, G., Vinson, D.P., Siri, S. & Cappa, S. (2003). An investigation of semantic errors in unimpaired and Alzheimer's speakers of Italian. *Cortex, 39,* 419-439.

Sera, M., Elieff, C., Forbes, J., Burch, M.C., Rodriguez, W., Dubois, D.P. (2002). When language affects cognition and when it does not: An analysis of grammatical gender and classification. *Journal of Experimental Psychology: General, 131*, 377-397.

Slobin, D.I. (1996). From "thought and language" to "thinking for speaking". In J. Gumperz & S. Levinson (Eds.), *Rethinking Linguistic Relativity*. Cambridge, MA: Cambridge University Press.

Vigliocco, G., Vinson, D., Indefrey, P., Levelt, W., & Hellwig, F. (2004). Role of grammatical gender and semantics in German word production. *Journal of Experimental Psychology: Learning, Memory & Cognition, 30*, 483-497.

Vigliocco, G., Vinson, D.P., Paganelli, F. & Dworzynski, K. (2005). Grammatical gender effects on cognition: Implications for language learning and language use. *Journal of Experimental Psychology: General*, 134, 501-520.

Whorf, B. (1956). *Language, Thought, and Reality: selected writing of Benjamin Lee Whorf*, ed. J.B. Carroll. Cambridge, MA: MIT Press.