ELSEVIER

Evidence against Whorfian effects in motion conceptualisation

Filippo-Enrico Cardini [*]

Lancaster University, Department of Linguistics and English Language, Lancaster, LA1 4YT, UK

Abstract

This research tested the linguistic relativity theory in relation to the conceptual domain of manner of motion. Nineteen English and 19 Italian native speakers completed two tasks involving the use of 26 triads of video-clips showing motion events. The participants underwent first a non-linguistic trial consisting of a forced-choice similarity judgement task performed during speech shadowing. Subsequently, they were asked to verbally describe the same stimulus material used for their similarity judgements. Congruently with the findings of Cardini (2008), an analysis of the verbal descriptions showed that English speakers provided much more information about the manner in which some motion occurs than Italian speakers. However, in contrast to the significant difference found across the two linguistic groups in the verbal task, the scores regarding the non-linguistic performances were close to identical: when visually attending to the motion events displayed in the video-clips, English and Italian speakers exhibited the same differential attention for manner vs. path of motion. The results of this study provide evidence against Whorfian effects on non-linguistic cognition.
© 2009 Elsevier B.V. All rights reserved.

Keywords: Linguistic relativity; Manner of motion; Language typology; English; Italian

1. Introduction

1.1. Whorf and linguistic relativity inquiries: do speakers of different languages view the world in a different fashion?

The idea that the particular language one speaks affects the way one thinks is traditionally associated with the theory of *linguistic relativity*, a term which was coined in the first half of the last century by Benjamin Lee Whorf (1897–1941):

> ... From this fact proceeds what I have called the ''linguistic relativity principle'', which means, in informal terms, that users of markedly different grammars are pointed by the grammars toward different types of observations and different evaluations of externally similar acts of observation, and hence are not equivalent as observers but must arrive at somewhat different views of the world (Whorf, 1956:221).

Whorf's often sweeping statements about language impact on thought, however, have never been easy to interpret, and so multiple versions of his theory ensued. There are *stronger* and *weaker* versions of his theory. The strongest formulation of the Whorf hypothesis, which views language as the only shaper and determiner of our ideas (a view that

---

* Present address: Piazza Bonavino, 2b/9, 16156 Genova, Italy. Tel.: +39 010 6983533.
  E-mail address: f_cardini_2000@hotmail.com.

leads to the extreme conclusion that no thoughts are possible without language), has now been dismissed by every researcher in the field as unsustainable. But some milder versions of the hypothesis are still taken in serious consideration by various linguists and psychologists. One can perhaps distinguish between at least two kinds of such milder versions. The weakest of the two concedes that language must have an impact on thought during but not beyond the time of speech, and it is often put in connection with the theory of so-called ''thinking for speaking'' (Slobin, 1987, 1996). The other version postulates language influence on thought even outside speech time. It is this second version which is going to be tested in the present study, since it seems to better address the rather stable and pervasive type of influence Whorf probably had in mind when postulating the capacity of a particular language to affect *habitual thought*. The Whorfian theory has been and is still fiercely opposed by many, especially by the so-called ''Nativists''. On their account, languages differ at the surface level but not at the deepest level of meaning, which is given by innate conceptual structures common to all human beings (Fodor, 1975; Jackendoff, 1983). Learning a language is simply finding the local phonetic clothing for pre-existing concepts: ''It is possible to suppose that these linguistic categories and structures are more or less straightforward mappings from a pre-existing conceptual space, programmed into our biological nature: humans invent words that label their concepts'' (Li and Gleitman, 2002:266). Empirical evidence, however, has never been really conclusive about the validity of the nativists' views, and controversies on the relativity issue are still far from being resolved. During the second half of the past century, many studies on linguistic relativity have often produced conflicting evidence. This has been the case with the domain of **colour**, for example, which has been widely tested in relation to possible *lexical influence* on thought. On the one hand, the influential study of Berlin and Kay (1969) demonstrated that we all perceive a common universal inventory of eleven basic colour categories, all or only some of which can be encoded by a language through 11 or fewer basic colour terms. Congruent with such findings were those of Rosch Heider's (1972) famous experiment on American vs. Dani people, which demonstrated a common underlying conception of colour relationships that is due to physiological rather than to linguistic constraints. On the other hand, however, the study of Davidoff et al. (1999) conducted between English vs. Berinmo speakers showed that different linguistic categorisations of the hue spectrum do produce non-trivial effects on colour perception. With regard to experiments geared at checking possible relativity effects from *grammatical influence*, again, no conclusive evidence seems to have been obtained by the various studies carried out in that respect. Positive evidence for linguistic relativity effects was found, for example, by Lucy (1992a), who tested English and Yucatec Maya speakers on the grammatical marking of **plurality**. Negative evidence was found, for example, by Papafragou et al. (2007) in a study contrasting English and Korean children on the domain of **evidentiality**. Other experiments such as that of Imai and Gentner (1997) conducted on English vs. Japanese speakers on the grammatical marking of **individuation** did find some relativity effects while also demonstrating, however, the universality of our non-linguistic cognitive distinctions between objects and substances.

The reasons for so much contrasting evidence on linguistic relativity phenomena are probably multiple. One issue surely regards the particular methodology adopted by the researchers in their experiments, which can often bias the results in one direction rather than in another. For example, in one particular test regarding memory of focal categories vs. non-focal, Rosch Heider's (1972) results matched those from Davidoff et al. (1999) when the experimenters used exactly the same procedure, but led to opposite conclusions when Davidoff et al. (1999) randomised the disposition of the colours array. Another issue concerns what particular aspect of thought is actually targeted by some linguistic relativity experiment. Gentner and Goldin-Meadow (2003:12) rightly argue that the term ''thought'' embraces a vast number of different cognitive areas and abilities: language may be able to affect some aspects of thought but not others. Boroditsky et al. (2003:63), for example, suggest that language influence is most powerful for more *abstract domains*, those not so reliant on sensory experience, whereas it may find considerable difficulty to emerge in domains heavily constrained by universals of physics and physiology. They may have come to such a conclusion after positive evidence for relativity effects was obtained both by Boroditsky (2001) in a study contrasting Mandarin and English speakers on the non-sensory domain of **time**, and by Phillips and Boroditsky (2003) in a study contrasting Spanish and German speakers on grammatical **gender**. (For those languages that attribute gender to nouns, the perceptual information available for most objects does not provide much evidence as to their linguistic gender, e.g. the visual perception of the sun does not suggest that the sun is masculine (e.g. Spanish) or feminine (e.g. German).)[1]

---

[1] With regard to the domain of time, however, note that a recent repeat of Boroditsky's (2001) experiments came to opposite conclusions (see January and Kako, 2007).

Their proposal, however, does not explain why, for example, some experiments carried out on the very much perceptual domain of colour provided positive evidence for relativity effects (e.g. those of Davidoff et al. (1999) mentioned before), nor does it explain, for example, why negative evidence was found by Papafragou et al. (2007) in their experiment on the domain of evidentiality, which, according to those researchers, is characterised by ''the subtleness and *abstractness* of its underlying concepts'' (Papafragou et al., 2007:253).

Despite such inconsistencies, one cannot exclude that the claim made by Boroditsky et al. may indeed have some overall kind of validity. If we take the domain of **space**, for example, those experiments which have focussed more on how people seem to *reason* about and organise the contents of what is perceived in space have produced encouraging relativistic results, whereas other experiments which have investigated aspects of spatial conceptualisation more directly tied to space perceptual apprehension have predominantly yielded evidence against relativity phenomena.

To the first kind probably belong the studies on **spatial frames of references**, which showed cognitive differences that result from differences in the linguistic coding of spatial relationships between entities, namely, from the use of absolute spatial terms (e.g. north/south/east/west) in languages such as Tzeltal, Arrernte, Guugu Yimithirr (see descriptions of such languages in Levinson, 2003), versus egocentric terms (front/back/right/left) in European languages, for example.[2]

To the second kind probably belong some past non-linguistic experiments made on the domain of **manner and path of motion**, in which speakers of two contrasting kinds of languages (the so-called ''satellite-framed languages'' and ''verb-framed languages'', to be discussed in the next section) were predicted to show a differential focus of attention on the manner vs. the path of some motion being visually attended to. As we shall see, the majority of the results from such experiments seems to suggest that visual processes are impervious to the workings of language, although evidence for linguistic relativity of this kind has been found in some cases.

The present study is aimed at gathering additional data precisely on this domain by contrasting English and Italian speakers on the particular dimension of manner of motion. With respect to the distinction made earlier between Whorfian effects occurring *during* speech time and those occurring *outside* speech time, it is important to note that only one of the past experiments carried out on the manner/path of motion domain has explicitly tested the latter type of effects (i.e. the experiment conducted by Gennari et al. (2002) in the ''Shadow'' condition). The contribution of the present study to the linguistic relativity research on motion conceptualisation is that it should provide further clues in relation to that stronger and more stable kind of influence, for which only a very small amount of empirical data is currently available.

In the remaining part of this introduction it will first be explained the different typology of expression that English and Italian speakers allegedly exhibit when describing motion events. It will then follow a short description of the methods usually employed by researchers for checking linguistic relativity effects in motion events conceptualisation.

1.2. Contrasting lexicalisation patterns for the expression of motion events

The way languages describe *motion events* has been a very popular topic in the recent literature of cognitive linguistics. The notion of such a kind of events is still largely put in connection with the theoretical work of Talmy (e.g. 1985, 1991, 2000). According to Talmy (2000:25), the description of a basic motion event is one in which some entity (the *Figure*) performs some translational *Motion* following a trajectory in space (the *Path*) in relation to some locative

---

[2] In brief, what was found in some non-verbal set of tests is that speakers of languages using absolute coordinates solved tasks about spatial arrangements in an opposite fashion to that shown by European speakers. In the so-called ''rotation paradigm'', for example, subjects saw an arrow on a table pointing to their left (for European speakers)/south (for speakers of absolute languages). They were then rotated 180° and asked to place the arrow on another table so it was just as before. If they pointed it to their left, they would think about it in terms of egocentric coordinates; if to their right (i.e. south), in geocentric coordinates. The results provided strong evidence for linguistic relativity: the large majority of subjects solved this non-linguistic task following the coding pattern of their language. While Europeans reversed the pointing direction of the arrow, speakers of absolute languages left it as it was originally. In relation to the point being made about conception vs. perception, one should observe that the objects in the visual field were certainly perceived in the same way by both groups of speakers. What was different was the orientation system organising the locative relationships between the entities perceived, something which, evidently, is not determined by our perceptual organs. Therefore, the linguistic influence found in this experiment must have involved a more abstract level of conceptualisation taking place upon the contents of the sensory input, a further processing of perceptual information that regards the way we think about, rather than apprehend, the space around us. It must be noted that Levinson's findings have been challenged by Li and Gleitman (2002) who concluded that we all think equally in relative or absolute frames of reference. Levinson et al. (2002), however, later responded to the challenge, reasserting the validity of their position.

reference (the *Ground*). To these four obligatory semantic elements, some optional ones such as *Manner* of motion, for example, can often be found. Two are, by and large, the fundamental combinations languages adopt to encode the semantic elements of a motion event onto syntactical and lexical structures. One first group of languages are called ''satellite-framed-languages'' (S-languages) because they typically encode Path through some verb particle attached to the main verb of the clause (a ''satellite'' to such a verb), whereas possible information about Manner of motion is conflated within the main verb's root. Typical S-languages are Germanic (e.g. English, German) and Slavic (e.g. Russian, Polish) languages, Finno-Ugric (Finnish, Hungarian), and Sino-Tibetan (e.g. Mandarin Chinese) languages. An example of this language type can be seen below in the following English sentence:

(1)   *Lisa* (F)   *skated* (Mo + Ma) *across* (P)   *the lake* (G)

A second group of languages are called ''verb-framed-languages'' (V-languages) because they typically conflate Path within the main verb of the clause; they leave possible information about Manner to be expressed at the end of that clause separately from the verb through an adjunct (an adverb, a gerundive form or a PP). Typical V-languages are Romance (e.g. Spanish, French, Italian), Semitic (e.g. Hebrew, Moroccan Arabic) and Turkic (Turkish) languages. An example of this language type can be seen below in the following Italian sentence:

(2)   *Lisa* (F)   *attraversò* (Mo + P)   *il lago* (G)   *con i pattini* (Ma)
         'Lisa   crossed   the lake   with the ice skates'

As one will have noticed, the construction exemplified in (1) allows a quicker expression of manner of motion than the construction exemplified in (2): because sentence (1) encodes manner in the main verb, there is no requirement of any adjunct at the end of the phrase, as it is the case in sentence (2) instead. Slobin (2003:162) observes that while for S-language speakers manner is ''for free'', for V-language speakers there is some ''cost'' involved in terms of further lexical production. For this reason, V-language speakers like Italian indicate manner of motion when it is contextually relevant, otherwise they do not mention it. One needs to say, however, that the bipartite typology of languages proposed by Talmy is not categorical. Berman and Slobin (1994:118, footnote 4) are aware that ''. . .categorical characterizations often reflect tendencies rather than absolute differences between languages''. That is, there is no absolute ban for any of the two kinds of languages to use the lexicalisation pattern opposite to that it is typically associated with. As one can certainly find instances of verb-framing in S-languages, so it is also possible to find instances of satellite-framing in V-languages (Aske, 1989), although it seems very rare to find V-language speakers use manner of motion verbs when describing motion events involving the *entering*, the *exiting* or *crossing* of something (so-called ''boundary-crossing-events'') (Slobin and Hoiting, 1994). But although Slobin (2004:26) has recently expressed the conviction that languages should be put on a continuum (a ''cline of manner salience'') rather than be forced either into the S- or into the V-group of languages, a fundamental divide between these two groups still holds on the grounds of empirical findings. For example, a considerable number of linguistic experiments involving descriptions of motion events have clearly showed that speakers of S-languages express manner of motion verbs much more frequently than V-language speakers (e.g. Naigles et al., 1998; Özçalışkan and Slobin, 1999; Papafragou et al., 2002; Cardini, 2008). The higher frequency with which manner of motion verbs are used in normal conversation by the former group of speakers has also indirectly emerged in some tests designed to measure how easily such verbs can be accessed. All experiments of this kind found that English speakers retrieve manner of motion verbs significantly more quickly than speakers of V-languages such as French (see Slobin, 2003:164), Basque (Ibarretxe-Antuñano, 2004:324), and Italian (Cardini, 2008:551–559). Further, it may not be a chance that the vocabulary of manner of motion verbs has been found to be significantly richer and more fine-grained in languages typically classified as S-languages than in languages typically classified as V-languages (Dukhovny and Kaushanskaya, unpublished results; Jovanović and Kentfield, unpublished results; Cardini, 2008).

1.3. Linguistic relativity effects in the domain of manner and path of motion?

The different degree of linguistic salience for manner of motion found between S- and V-language speakers raises the question of whether or not this difference can also be found at non-linguistic level because of Whorfian effects. The method normally adopted for answering this question involves the use of triads of pictures or video-clips, each of them portraying the occurrence of some motion event.

One type of task carried out using such triads is that in which participants are asked to make *similarity judgements* between the three clips of one triad. Each triad is composed of one *target* clip/picture and two other clips/pictures similar to it, the so-called *variants* or *alternates*. The target item shows an entity performing a motion that has a certain manner and path (e.g. a ball rolling into a pipe); one of the two similar variants is identical to the target item except for the path of motion (e.g. the ball rolling out of the pipe), whereas the other is identical to the target except for the manner of motion (e.g. the ball bouncing into the pipe). The participants are required to indicate which variant is more similar to the target. If they choose the same-manner variant, the dimension of manner is thought to be more salient to them than that of path because it is the one determining the idea of greater similarity to the target. The opposite is true if they choose the same-path variant. The usual line of reasoning is that, because S-language speakers express manner more often than V-language speakers, on average, the former should prefer same-manner variants significantly more often than the latter.[3] Previous studies which have involved some similarity judgement task on the domain of manner and path of motion have produced conflicting results, although the overall bulk of evidence seems to support the view of linguistic relativity sceptics. Negative evidence for linguistic relativity in similarity judgment tasks was found in the studies of Bohnemeyer et al. (2001), of Papafragou et al. (2002), of Gennari et al. (2002) (in the ''Free naming'' and in the ''Shadow'' condition), and of Pourcel (2005). Positive evidence was found in the studies of Gennari et al. (2002) (in the ''First naming'' condition), of Hohenstein (2005), and in Experiment 1 of Finkbeiner et al. (2002).[4]

Another kind of task carried out in the past using triads of pictures or of video-clips has sought to find possible influence on *memory*. In the *recognition memory* task, for example, participants are first shown a certain number of target clips, one at a time. Subsequently, in a second session of the same task, they are shown in a random order both the target items viewed in the first session and their variants (same-manner and same-path variants); for each video-clip they must say whether or not they have seen that clip in the first session of the task. The degree of manner vs. path salience displayed by participants emerges from the differential distribution of recognition *errors* they make across the two variants of each triad. For example, because language-based predictions expect V-language speakers to focus their attention more heavily on path rather than on manner, they should show a higher ratio of recognition errors on same-path variants relative to same-manner variants than S-language speakers; by contrast, because the latter are expected to focus their attention more heavily on manner rather than on path, they should display the opposite pattern of recognition errors. Hitherto, recognition memory tests conducted on S-language vs. V-language speakers (Papafragou et al., 2002; Gennari et al., 2002) have produced negative evidence for relativity phenomena. The only experiment which did find some linguistic influence on memory is that reported in Pourcel (2005). This experiment did not employ triads of video-clips, but consisted in first showing participants a four-and-a-half-minutes film containing many manners and paths of motion, and then in having participants fill out a questionnaire which would check the accuracy with which the dimension of manner and path had been attended to.

2. Description of the study

2.1. Planning the experiment

2.1.1. Choosing the type of task

As was just said in section 1.3, there are, by and large, two kinds of non-linguistic task based on event triads that have been conducted so far: *similarity judgement tasks* and *recognition memory tasks*. When comparing the two kinds of task, recognition memory probably involves more subconscious processes than similarity judgements. Gennari et al. (2002:56) argues that ''recognising an event as previously stored in memory...can be carried out automatically by

---

[3] But Papafragou et al. (2002:198) also propose a second possibility which runs in the opposite direction. Because the verb is the ''informationally privileged element'' and in manner-verb languages (i.e. S-languages) path is exhibited on the surface of the motion verb sentence (e.g. *across* the road), while manner is ''hidden'' inside the meaning of the verb, and the opposite is true for path-languages (i.e. V-languages), speakers of S-languages should actually be more sensitive to path, while speakers of V-languages to manner. As a consequence, the former should be expected to choose same-path variants more often than the latter.

[4] Note that Finkbeiner et al. regard their study as one on memory rather than one on similarity (indeed the title of the study is *The Role of Language in Memory for Actions*). However, both Experiment 1 and Experiment 2 of their study actually involved a similarity judgement task of the same kind as that already outlined in this section. The researchers classified their study as one on memory because, as we shall see in section 2.1.3, they found some evidence for the fact that in Experiment 1 informants used inner speech as a means to enhance memory of the target item.

retrieval processes that give a sense of familiarity, and can be accomplished in the absence of conscious awareness...". By contrast, "judging similarity in the absence of time constraints depends on aspects of the comparison process that may emerge after some deliberation" (Gennari et al., 2002:56). Since one of the Whorfian claims is that relativity effects take place especially at subconscious level (Lucy, 1992b:46), the recognition memory task should therefore be more revealing than that of similarity judgement as a means of investigation. Other considerations, however, seem to raise some doubts about the reliability of the results that could come out of a recognition memory task.

One such consideration is that the recognition memory task requires an adequate degree of difficulty for maximising the chances of detecting possible linguistic effects, and to find this appropriate level of difficulty may not be easy at all. In the previous section we saw that the degree of manner vs. path salience emerges from the differential distribution of errors made by the two groups across the two variants of one triad. Differently from similarity judgement, where *all* responses are informative of manner vs. path salience (and where no response is either correct or incorrect), here only incorrect answers are so, because it is from their patterns that one can infer the presence of linguistic influence. Therefore, it is important that the task generates a sufficient number of errors: if the task is too easy and thus the number of errors is extremely low for both groups, the different patterns of errors made by the two groups will probably not be able to emerge in any significant way. Likewise, one should make sure that the task is not too hard either, or else both linguistic groups will produce an extremely high and probably similar number of errors, whereby, again, the differential sensitivity to manner and path between English and Italian speakers might not be able to come to surface. Getting an adequate degree of difficulty for the memory task is therefore an important issue for recognition memory. But, as already said, this may be hard to accomplish. The difficulty of the task depends on many factors, like, for example, the number of triads used in the experiment and the length of the possible distracting activity the researcher may want to set between the viewing of the target clips and that of all clips mixed together (targets and variants). It is virtually impossible to know in advance what the most adequate values for each of these variables are.[5] Similarity judgement, on the contrary, does not pose any concern in this respect, since it does not need to generate a suitable number of errors, a notion which, anyway, is alien to this kind of task altogether.

The other problem with recognition memory might arise from the possible repetition of some very common manners (e.g. *walk*, *run*, *jump*), or even just the presence of similar manners (e.g. different kinds of *jumping*) across different triads. In recognition memory, this could be a confusing factor when participants try to remember whether they have already seen what they are watching. But in similarity judgement this is not an issue: the response is based on what is currently being watched, and the triads watched previously are very unlikely to exert any influence on each decision being taken.

For these reasons it was decided that the relativity investigation carried out here should only involve a similarity judgement task.

2.1.2. Issues concerning the stimulus material

One of the main concerns about the testing material regards *the nature of the stimulus*. Since the object of the present enquiry involves motion events, and since the central character of motion is dynamicity, the stimulus to be presented to the informants has to have dynamic character, and not consist of static pictures as in Papafragou et al. (2002), for example. In other words, it has to be a kind of stimulus in which movements can actually be perceived rather than inferred. Therefore, as in Bohnemeyer et al. (2001), Finkbeiner et al. (2002), Hohenstein (2005), Gennari et al. (2002), and Pourcel (2005), the stimulus material used in the present study will consist of video-clips.

As for *the contents of the clips*, like all previous similarity judgement tasks on manner of motion, each video-clip has to portray some entity performing a translational motion with some manner through some path. Video-clips will be grouped in triads, and the clips of one triad will relate to each other in the way illustrated earlier in section 1.3. It must

---

[5] The importance of the variable concerning an adequate interval time between encoding and recall in relation to the difficulty of the memory task, for example, was recognised by Pourcel (2005) in the "late recognition recall" (session 2 of the memory experiment). Because the one day interval between encoding and recall was felt to be rather long (participants from both groups would have probably remembered too little of the film for possible relativistic effects to be able to emerge), the experimenter decided to have participants view the film a second time with a live commentary of it after the end of session 1, and to re-tell the story before the start of session 2. Evidently, given that that recognition recall did yield some evidence for relativity effects, the balance struck between the task difficulty and interval time separating encoding from recall was sufficiently adequate in relation to the goal of the test. Nevertheless, no rigorous empirical method could be employed for determining the adequacy of that balance, and the good judgement of the experimenter on that may have been simply fortunate coincidence.

be pointed out, however, that in this study the dimension of path will be used purely as a control for measuring manner salience. That is to say, if linguistic group A makes, when compared to linguistic group B, a much higher number of same-path choices, this should not be interpreted as group A actually showing a higher degree of salience for the path dimension, but rather a lesser degree of salience for the manner dimension (which makes speakers of group A opt for the same-path variant more often than those of group B). The reason for this approach is that, in purely theoretical terms, S- and V-languages are thought to differ in the amount of linguistically expressed manner (i.e. the former express it more often), but not in the amount of path. This is because, in basic motion event descriptions the presence of Path, as well as that of Figure, Motion and Ground, is obligatory (see section 1.2). Thus, both S- and V-languages should not differ in the attention linguistically given to path, and no linguistic effects on thought should therefore arise as a consequence for this dimension. Because path is here intended to be used simply as a control for manner focus, in some of the triads the supposed different path showed by the same-manner variant will not actually be a different path, but simply the opposite direction of one and the same path (e.g. same path *across*, whereby the target clip shows an across-motion going from left to right, and the same-manner variant shows the across-motion going from right to left). Like two actually different paths, two opposite directions of one same path are just as able to provide the control factor needed in order to measure focus on manner.

Still with regard to the content of the clips, my sample will not include any instance of events involving "caused motion" (e.g. see Pourcel, 2009). A typical caused-motion-event is one in which the motion of some Fig. 1, which is the motion expressed by the verb of the clause, triggers the motion of some Fig. 2. An example of caused-motion-event can be "He (Fig. 1) kicked (Mot.) the ball (Fig. 2) across (Path) the street (Ground)". Although the verb expressing the motion of Fig. 1 is almost certainly loaded with some manner (i.e. the particular movement of the leg while being in the act of kicking), such a manner clearly does not refer to a "basic motion event" as conceived by Talmy, that is, as a translational kind of motion *following a path* in space (Talmy, 2000:25). In the above example, this kind of motion is not performed by Fig. 1, but by Fig. 2 (the ball) in its crossing of the street. However, we do not know anything about the manner of this motion: is the ball *flying*, *bouncing* or *rolling*? Caused-motion-events are often included in the list of event triads employed by linguistic relativity studies on the domain of manner and path of motion (e.g. Gennari et al., 2002). However, I believe that to provide samples of triads showing caused- and non-caused-motion-events mixed up together is methodologically questionable: the kinds of conceptualisation involved in these two types of event are clearly different and, in the present view, not compatible with one another.

With regard to *the size of the sample* to be shown to the participants, the number of manners should be fairly large. Only in this way will the sample be able to be sufficiently representative of the domain investigated. In this specific respect, a study such as that of Gennari et al. (2002), for example, which involved a very large number of manners and paths is certainly to be preferred to one such as that of Bohnemeyer et al. (2001) which offered only four different manners and two different paths.

As for *the time-length of the clips*, this should be fairly short in order to minimise the risk that the participants start to reason and reflect upon the contents of the clips, and, perhaps, even end up adopting some conscious strategy for answering once they have understood the pattern of clips presented to them. In particular, once some participant has realised that one of the two variants always holds the manner of motion unvaried, whereas the other holds the path/direction of motion unvaried, s/he might decide to always choose the same kind of variant in order to be consistent. In this sense, some concerns should be raised about similarity judgement tasks which do not force participants into rapid, immediate decisions due to the lack of time constraints on the viewing of the stimuli (as apparently was the case in Papafragou et al., 2002, for example), thus potentially favouring the adoption of a fixed pattern of response.

One consideration also goes to *the order of presentation of video-clips* within each triad. In the present investigation, this will be the same as that usually implemented in studies involving event triads: the target clip will be played first, and, once it has ended, the two variants will appear and be played side by side simultaneously. A different ordering system was employed in Pourcel (2005), in which the three items were shown one at a time with the participant being free to choose any two items of the triad which s/he perceived to be more similar out of the three. The problem with this order of presentation is that the similarity choice may be affected by the different lengths of time separating two different pairs of items (i.e. *item 1 ↔ item 2* as opposed to *item 1 ↔ item 3*), and also by the different length of time elapsing between the final participant judgement and two different pairs of items (*time of judgement ↔ pair 1–2* as opposed to *time of judgement ↔ pair 2–3*). In other words, short-term memory performances not being equal across different paring options, it cannot be excluded that some pairing

options are more or less likely to be made than others because of the influence of those differential memory performances.[6] Another different order of presentation was employed by Finkbeiner et al. (2002) with some participants of Experiment 2, where the target and the variants of each triad were shown simultaneously. Some concerns about this ordering system may arise in relation to the issue of covert speech, which will be discussed in the next section.

2.1.3. Shadow condition

According to Lucy (1996:48; 2000:xii), an experiment bears on linguistic relativity when it does not involve any verbal task. The belief behind this claim is that linguistic relativity inquiries should investigate a *permanent* rather than a temporary linguistic influence on thought. Language temporary influence on various aspects of cognition is actually an established fact,[7] and so to have participants produce or experience language during or immediately prior to some cognitive task would preclude the inquiry into a more stable and long lasting kind of influence. Usually, similarity judgement tasks do not involve any explicit linguistic activity, since the participant is simply asked to indicate the preferred variant (e.g. by pointing with a finger). However, during the viewing of the target clip, the participant could well produce some inner speech, which could ''invalidate'' the linguistic relativity status of the experiment. For example, during the watching of a target clip showing a man running into a tunnel, the participant could produce the verb *run* covertly and then choose the same-manner variant because of the *temporary* influence of that linguistic concept activated an instant earlier. This possibility is not too remote: Experiment 2 of Finkbeiner's et al. (2002) study provided some evidence for the fact that the linguistic effects found in Experiment 1 might indeed have resulted from participants making use of inner speech; Hohenstein (2005:420–421) herself is aware that the positive evidence she found for linguistic relativity may have been brought about by covert language, and suggests that the same experiment could be replicated with participants being asked to perform a distracter task.

One way of preventing participants from engaging in covert language could be that adopted by Finkbeiner et al. (2002) with some participants of Experiment 2, where all three items of a triad were presented on the screen simultaneously. The overall results obtained in that study seem to suggest that covert language is only employed as a means for memory enhancement (e.g. in similarity judgement tasks where the two variants appear on the screen after the target has vanished, so that the latter has to be ''remembered''); thus, for tasks where no memorisation strategies are required, such as similarity judgement tasks with all three items of a triad shown at the same time, instances of inner speech should not take place. However, although such a conclusion does rest on some empirical evidence collected in that study, the method of simultaneously presenting all items of a triad to the participant might still be regarded as not sufficiently safe. It may well be that covert language occurs outside voluntary control, and not always with the specific aim of helping memorisation. Inner verbalisation seems to be just as likely to spontaneously occur in absentminded daydreaming as in various kinds of problem solving, and without us being fully aware of that (Hurlburt, 1990). In which case, the occurrence of inner speech during the simultaneous presentation of the three items cannot be excluded. Although, of course, one cannot definitely produce three different linguistic descriptions at the same time, the possibility of some participant covertly describing only one of the three items (possibly the target item, which stands out from the two variants because of its more separate position on the screen) cannot be entirely ruled out.

A safer measure to prevent participants from engaging in covert language should be that of running the relativity experiment in a *shadow condition*. The shadowing consists in having the participants repeat aloud a series of random numbers or letters played by a tape recorder while they undergo the similarity judgement task. Such an activity loads their working-memory, so that further linguistic production is impeded, or at least minimised. In relation to what was said in the previous section about the risk of participants engaging in conscious response strategies, it is important to

---

[6] Pourcel (2005) was aware of the problem, since she decided to re-order the videos of one of her triads in order to check for possible order effects (triad # 9 was the same as triad # 3, but with a different sequence of the three videos composing it). And when one looks at the results related to that particular pair of triads only, with the exception of the responses shown by the English group in Experiment 1, a non-trivial difference can actually be found between those two triads in the performance of both linguistic groups (Pourcel, 2005:223–224, 231); the difference shown by the French group in Experiment 1 is particularly significant.

[7] Already Carmichael et al. (1932) showed that the verbal labelling of some figure can bias the successive recalling of its shape. Bower et al. (1975) provided evidence for the fact that the recall of pictures is positively affected when verbal cues were given during memory encoding. Also some past studies on manner and path of motion have produced evidence for transitory linguistic influence on thought. For example, Gennari et al. (2002) found some temporary language bias in the ''Naming-first condition'' of their similarity judgment trial, in which participants had to describe the contents of some clips while performing their non-linguistic task on those clips.

remark that the shadowing activity should also provide a fairly effective measure against that potential risk: having to perform two tasks at the same time (audio-motor + visual) presumably leaves little room to the participant for reflecting about what is being viewed. The shadow condition should therefore also favour a kind of response based on the immediate sensation of the participant, that is, the kind of rather subconscious response we seek.

2.1.4. Languages investigated and correlated linguistic task

The languages that will be compared in the present study are English (an S-language) and Italian (a V-language). In a recent study (Cardini, 2008), these two languages have clearly shown to differ significantly in the degree of linguistic salience of manner of motion. For example, English speakers exhibited a greater frequency and variety of use of manner of motion verbs in their spontaneous speech, and were also significantly quicker in retrieving such verbs from memory. As a consequence, to also contrast these two languages at non-linguistic level should be a reasonable choice for exploring linguistic relativity phenomena. Given that evidence for a significantly different degree of linguistic salience of manner of motion between English and Italian speakers was already provided by Cardini's (2008) study, one might argue that no further check is actually needed in that respect. In the present investigation, however, on the same stimulus material used for the non-linguistic task a correlated linguistic task will also be conducted. This is because the results of some linguistic relativity experiment will certainly sound more convincing if non-verbal performances can be directly correlated with verbal performances associated with the same specific stimulus rather than correlated simply with general patterns of linguistic behaviour. In this respect, the conclusions reached in studies in which some non-verbal task was not paralleled by a correspondent verbal task (e.g. Bohnemeyer et al., 2001; Finkbeiner et al., 2002) are simply based on a *presumption* made on the linguistic side, namely that if the speakers of the languages investigated had also been tested linguistically on the particular material selected for their non-linguistic task, they would have certainly shown a significantly different performance, and in the way predicted by Talmy's categorisation.

One important thing to ensure is that the linguistic task must take place *after* the non-linguistic task. The reason for this has to do again with the transitory linguistic influence that the verbal activity could exert upon the non-linguistic task performance if conducted first. In this case, at the time of the non-linguistic task, the participant could still be under the influence of the linguistic encodings made in the verbal performance; and because language-based predictions expect English participants to produce more linguistic information about manner of motion than the Italian participants, they could be biased into choosing the same-manner option more often than the latter.

One should also observe that it will be sufficient to run the linguistic task on target clips only. In the initial non-linguistic task, the participant's choice of most similar variant will be based on how *the target clip* has been encoded non-linguistically (e.g. if his/her non-linguistic encoding has predominantly focused on manner of motion, s/he will probably pick the same-manner variant, and vice-versa). Therefore, to investigate Whorfian effects, it will suffice to see whether there is any correspondence between the linguistic and the non-linguistic encoding of the target clips (whereby the latter encoding will be revealed by the choices made on the variants).

As for the instructions necessary for the task, these should not impose too many constraints on the desired description. It is important that the participants are left free to come up with a rather instinctive utterance without first being obliged to think too much about what they should/should not say. Only in this way will the elicited linguistic sample be able to reflect patterns of expression habitually used in every day conversation. In this respect, the instructions given by Gennari et al. (2002), for example, which involved illustrative answers showing both the desired structure of the phrase (to the length that they even indicated in parentheses possible optional forms for the Spanish speakers), and the undesired one, may have inhibited an immediate, spontaneous response. To ask participants to simply produce a short utterance describing the main action shown in the clip is perhaps a simpler and better way to look into habitual, subconscious patterns of expression.

2.2. Method

2.2.1. Participants

19 monolingual English speakers, undergraduates at the University of East Anglia (UK), and 19 monolingual Italian speakers, undergraduates at the University of Genoa (Italy), participated to the test. The English group was composed of 10 male and 9 female participants; the Italian group of 9 male and 10 female participants. The average age of the English participants was 19.4, that of the Italian participants was 22.3. In order to ensure the best possible

homogeneity of the sample, all 38 participants were recruited from the Faculty of Humanities. The recruitment procedure involved potential participants filling in a questionnaire in which they had to give information about their native language and about possible knowledge of foreign languages. Only those English and Italian volunteers who declared not to be fluent in any foreign language were selected for the experiment.

2.2.2. Material

The testing material consisted of 26 triads of video-clips, all of them showing some entity performing a motion event. The way the clips composing one triad related to each other was the one already described at the beginning of section 1.3. The average duration of a clip was 3.6 s. The interval between each target clip and the two variant clips was 1.5 s. The three clips of a triad had an identical number of frames, thus enabling the synchronous presentation of the two variants. The contents of the video-clips did not involve ''artificial animations'', but were ''natural'' in that they showed real people or objects performing mannered motion events in real life environments. In relation to the manner element, the element specifically targeted in this investigation, the sample portrayed, overall, 25 different manners of motion which can evoke either the manner verbs reported below in Table 1 or similar verbs. Further, the sample was meant to portray 12 different paths of motion; these are also to be found below in Table 1.

Table 1
List of triads used in the experiment (the > and < symbols stand for two contrasting directions of one same path).

	Target	Same-manner	Same-path
1	Man struts towards tree	Man struts away from tree	Man walks towards tree
2	Man walks up stairs	Man walks down stairs	Man runs up stairs
3	Ball rolls into pipe	Ball rolls out of pipe	Ball bounces into pipe
4	Man runs across corridor	Man runs along corridor	Man stumbles across corridor
5	Girl cycles around > car	Girl cycles around < car	Girl rollerskates around > car
6	Man climbs down a slide	Man climbs up a slide	Man slides down a slide
7	Man tiptoes into building	Man tiptoes out of building	Man sneaks into building
8	Man strides through chairs	Man strides around chairs	Man shuffles through chairs
9	Man skips around tree	Man skips away from tree	Man runs around tree
10	Man leaps across > road	Man leaps across < road	Man jumps across > road
11	Man bounds into tunnel	Man bounds out of tunnel	Man jumps into tunnel
12	Man jumps towards tree	Man jumps around tree	Man skips towards tree
13	Man marches across > road	Man marches across < road	Man strides across > road
14	Boy crawls out of car	Boy crawls into car	Boy steps out of car
15	Ball bounces through net	Ball bounces over net	Ball rolls through net
16	Girl runs around > chair	Girl runs around < chair	Girl rollerskates around > chair
17	Man stumbles across > corridor	Man stumbles across < corridor	Man shuffles across > corridor
18	Man climbs up ladder	Man climbs down ladder	Man jumps up ladder
19	Man marches towards car	Man marches away from car	Man limps towards car
20	Man skips through chairs	Man skips around chairs	Man hops through chairs
21	Man jumps around ladder	Man jumps up ladder	Man jogs around ladder
22	Man staggers away from tree	Man staggers towards tree	Man marches away from tree
23	Man walks across stairs	Man walks up stairs	Man strides across stairs
24	Can rolls across > table	Can rolls across < table	Can slides across > table
25	Man bounds out of tunnel	Man bounds into tunnel	Man leaps out of tunnel
26	Boy jogs around table	Boy jogs past table	Boy hops around table

To give a practical example of triad, Fig. 1 shows pictures from triad # 5 (see next page).

2.2.3. Tasks and procedure

Each English participant was tested individually in a room made available by the University of East Anglia; each Italian participant was tested individually in a room made available by the University of Genoa. In order to minimise stress levels, participants were reassured that none of the tasks they were about to absolve involved correct or incorrect answers, so there was no such a thing as ''good'' or ''bad'' responses. The participants sat in front of the lap-top computer, and the researcher sat beside them. The overall duration of the testing session for each participant was about 20/25 min.

Fig. 1. Pictures from triad # 5 (target: on top; same-manner: bottom left; same-path: bottom right).

The participants engaged in the non-linguistic task first. They were presented with a sheet of paper with the task instructions written in their native language. Below is the English version:

> ''In this task, you will be presented triads of video clips on the laptop screen. For each triad, you will first see one only clip appear and be played. Once finished, this will go off the screen, and the other two clips of the triad will simultaneously appear and be played together side by side. You will have to point with your finger which of these two clips, in your opinion, is more similar to the one previously viewed''.

Before the test started, participants were given two practice items so that they could become familiar with the task. In the first practice item, the target scene showed a man doing press-ups inside a room; one of the two variants showed the same man doing sit-ups inside the same room; the other variant showed the same man doing press-ups outdoors. In the second practice item, the target scene showed a man writing while seated at a desk in front of a pc; one of the variants showed the same man typing on the keyboard of the pc while being seated in the same spot as before; the other variant showed the same man writing on a different desk placed in another room with a different background. The 26 triads were presented in two orders (A and B), one the inverse of the other: for each linguistic group, 10 participants followed order A; the other nine followed order B. For each triad, the side on which the same-manner and the same-path variants appeared on the screen was randomly assigned. The same-manner variants appeared 14 times on the right side of the screen and 12 times on the left side. The triads were stored as individual files on the experimenter's PC laptop, and the experimenter started the presentation of each triad with a mouse-click when participants were ready. Before the test started, participants were also instructed to repeat aloud a series of numbers (ranging from 0 to 10) that would be randomly played by a cassette-player during the task. The rate at which numbers were played was about one per second. The researcher recorded the participants' choices on a score sheet.

Once the non-linguistic task was over, the participants engaged in the verbal task. For the reasons explained in section 2.1.4, this task was carried out on the target items only. Below, one can see the English version of the given instructions:

"In this task you will be presented some of the clips previously viewed, one at a time. After having watched one clip, give a verbal description of what happens in it. Just say one short sentence that simply focuses on the main action going on in the clip, and do not digress in details. The short sentence should answer the question: "What happens in the clip?"

All participants' responses were audio recorded on a tape and orthographically transcribed.

2.2.4. Results and discussion (verbal task)

Before looking at the results concerning the non-linguistic experiment, it is important to see whether a significant linguistic difference was actually found between the two linguistic groups, so that Whorfian effects possibly found in the non-linguistic test could indeed be explained in terms of language influence. Thus, the results of the verbal task will be reported first.

The results confirmed predictions made on the basis of the previous linguistic experiments reported in Cardini (2008), that is, English speakers gave an indication of the manner in which a motion occurred significantly more often than Italian speakers. If one looks at *all kinds of information given about manner of motion* (i.e. information contained either in the main verb, or in some adverb, some gerundive form, some PP, etc.), English participants gave such kind of information in 96% of their utterances, whereas Italians gave such information in 63% of theirs. A chi-square test indicates that such a difference is extremely significant ($\chi^2 = 32.57$; $p < 0.0001$). The results are illustrated in Fig. 2 below.
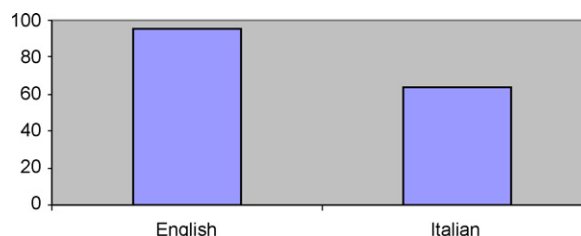
Fig. 2. Verbal task: percentage of utterances with manner information.

Also with regard to the number of utterances containing further elaborations of manner information, that is, utterances in which a main manner verb was coupled by some manner adverb, etc. (e.g. "A man is *walking slowly* towards the car"), this was higher in English participants (39 against 27 produced by the Italian participants).

But the difference between the two groups is even more significant if one considers only *manner information given by the main verb* (English = 93 %; Italian = 45%).[8] In this respect, results are similar to those obtained by Naigles et al. (1998) in a study between English and Spanish speakers involving the viewing of pictures (English used manner verbs in 91% of the cases, whereas Spanish in 47% of the cases), and to those obtained by Pourcel (2005) for English vs. French involving the viewing of video-clips (English used manner verbs in 85% of the cases, whereas French in 33% of the cases). Also remarkable is the difference found between the two groups with regard to the number of manner verb types, which was 33 for English and 15 for Italian. The results concerning manner information delivered by the main verb are illustrated in Fig. 3 below.

---

[8] The results regarding this second kind of information refer to 24 triads only, since they exclude the data collected for triad # 5 and # 6. This is because the target items of those two triads could not have elicited, and indeed did not elicit, any manner of motion verb from the Italian speakers. Italian does not have any manner of motion verb for describing the motion evoked by the English *cycle/bike* (in target item # 5), nor does it have any manner of motion verb that can be used to convey the concept of the English *climb down* (in target item # 6). By contrast, for all the remaining 24 target items, Italian speakers did have and indeed used manner of motion verbs for describing the motions shown in such items, even though, as expected, their verbs were often less fine-grained than those used by the English speakers. For example, Italians had to rely systematically on the general verbs *saltare/saltellare* (labelling the motion of *jumping* in general) to describe different particular kinds of jumping such as *hop*, *leap*, *skip*, *bound*, *spring*, etc.; likewise, they had to rely on *camminare* ('walk') to describe different particular kinds of walking (*stride*, *tiptoe*, *strut*). To ensure a balanced comparison between the performances of the two groups, also the numbers concerning the distribution of manner information across the average sentence (later to be found in Fig. 4a and b) and those concerning the kinds of verb used as main verb of the sentence (later to be found in Fig. 5a and b) will exclude triads # 5 and # 6.
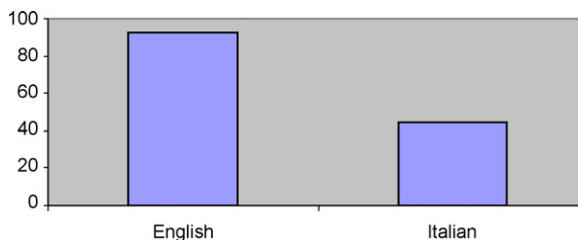
Fig. 3. Verbal task: percentage of utterances with manner information in main verb.

The quantity and distribution of manner information in the average sentence is summed up in Fig. 4a and b.
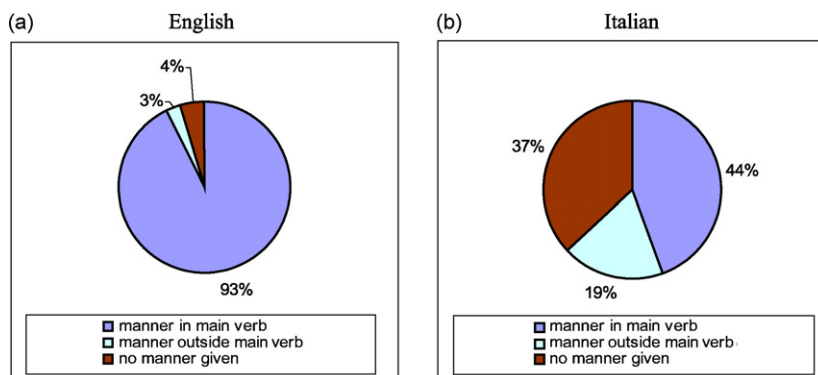

Fig. 4. (a and b) Verbal task: quantity and distribution of manner information in the average sentence.

The figures clearly show that English speakers gave manner information almost systematically through the main verb of the clause, a solution which appeared to be adopted "by default". As a result, there were very few isolated cases of manner information provided in some other way. By contrast, the number of times in which Italian speakers put manner information outside the main verb was not at all irrelevant, as one can see in Fig. 4b. As already established in Cardini (2008), it seems that Italian manner verbs are indeed less salient than the English ones, so that alternatives to expressing manner semantics via the main verb are sometimes sought. Italian path verbs, by contrast, seem to be much more salient than the English ones, as an analysis of the kinds of verb used as main verb of the clause demonstrates (Fig. 5a and b below).
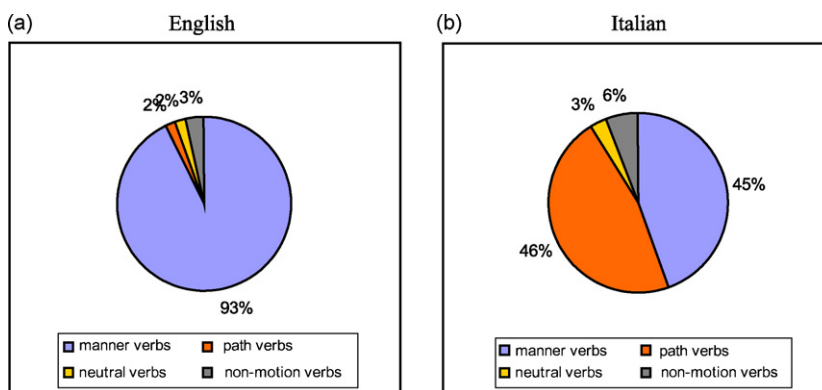
Fig. 5. (a and b) Verbal task: quantity of different kinds of verb used as main verb of the clause.[9]

---

[9] The label "neutral verbs" to be found in Fig. 5a and b refers to motion verbs such as the English *go*, *come*, *move*, which do not indicate either the manner or the path of translational motion. The label "non-motion verbs", by contrast, refers to verbs such as *make* or *take* which do not have any direct motion semantics, and can, at best, only imply some; some participants used these verbs in expressions such as "a boy *makes* big jumps across the road" or "a man *takes* long strides through some chairs".

It is also interesting to note that, when looking separately at the results concerning elicited boundary-crossing events vs. elicited non-boundary-crossing events, the difference in quantity of manner of motion verbs produced by the two groups (with English participants always giving more manner information than Italian participants) was more pronounced in boundary-crossing events. The English rate of manner of motion verbs per utterance remained unvaried irrespective of whether a scene had been conceptualised as a boundary-crossing-event (92%) or as a non-boundary-crossing event (93%). Italian speakers, however, clearly emerged as having more problems in using a manner of motion verb in scenes that they conceptualised as boundary-crossing-events (12%) than in scenes that they conceptualised as non-boundary-crossing events (51%). This is in line with those theoretical observations according to which V-languages have a severe difficulty in using manner of motion verbs only in boundary-crossing event descriptions (see section 1.2).[10]

2.2.5. Results and discussion (non-verbal task)

Differently from the verbal task, no significant difference was found between the two linguistic groups in the non-verbal task. The percentage of same-manner variants chosen by the English participants as most similar to the targets was nearly identical to that displayed by the Italian participants (44.9% for the former; 45.3% for the latter). A chi-square test confirms that such a difference is not significant ($\chi^2 = 0.009$; $p = 0.9246$). These results can be viewed in Fig. 6.
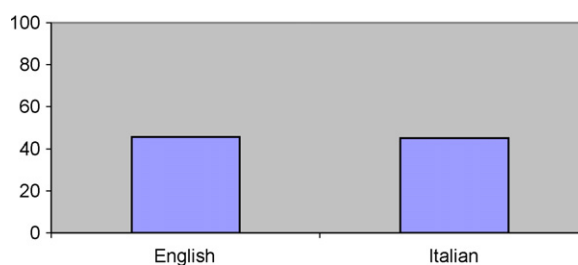
Fig. 6. Percentage of same-manner choices.

Obviously, the insignificant difference of same-manner preferences across the two groups provides evidence against linguistic relativity irrespective of what kind of language-based prediction one subscribes to: to the traditional one expecting V-language speakers to have the manner of motion less salient than S-language speakers, or to the opposite one suggested by Papafragou et al. (2002) (see again footnote 3).

The difference between the two groups remains insignificant if one considers only those triads for which the linguistic difference found between English and Italian participants in relation to quantity of manner of motion mention was greatest. In this respect, the divergence between the two groups seems to remain insignificant irrespective of the number of triads one wants to analyse which are most relevant in this sense. A check made only on those nine triads where the linguistic difference between the two groups was greatest in terms of quantity of manner information delivered (i.e. triads # 2, 3, 6, 8, 15, 13, 17, 18, 23) indicated that the difference is still not significant ($\chi^2 = 0.103$; $p = 0.7488$); another check made only on the five most relevant triads (i.e. triads # 2, 6, 15, 18, 23) came once again to the same conclusion ($\chi^2 = 0.099$; $p = 0.7532$).

The results of this experiment therefore provide clear evidence *against* linguistic influence on thought, although, of course, one cannot rule out that linguistic relativity effects were not detected because of the particular methodology adopted. In which case we should ask ourselves what features of the test may have prevented us from finding any linguistic influence.

The first possibility relates to the domain of manner of motion being one that is heavily constrained by universal mechanisms of visual perception. Because of such heavy constraints, linguistic influence might affect this conceptual domain only weakly, so that only particularly sophisticated stimulus material can bring it to light. For instance, it may be that some of the triads were constructed in a way which was not sufficiently adequate to prevent universal

---

[10] The analysis of manner of motion verbs quantity in relation to elicited boundary vs. non-boundary motion *events* necessarily excluded manner of motion verbs found in elicited motion *activities* (Pourcel, 2009), that is, motions which lack the path element (e.g. ''a man is walking''; ''a can is rolling on a table'').

perceptual mechanisms from biasing *all* participants too strongly towards one of the two choices, so that no significant difference between the two groups could possibly be found. To give an example of what is meant by this, in one particular triad used for the experiment (i.e. triad # 23), 35 out of the 38 participants opted for one and the same variant. The target clip of that triad showed *a man walking across some stairs*. The same-manner variant showed *the same man walking up those stairs*, while the same-path variant showed *the same man striding across those stairs*. The difference between the two manners involved in the triad was probably very small in relation to the difference between the two paths. While *stride* can be viewed simply as a kind of *walk* and is therefore very similar, *across* and *up* are rather different: the former is usually associated with the horizontal dimension, the second with the vertical dimension. Also, the target clip and the same-path variant both showed the moving man side-on, whereas the same-manner variant showed him with his back to the camera. In purely perceptual terms, the combination of at least these two factors must have strongly biased the choice of both English and Italian participants towards the same-path variant, rather than otherwise. And the bias may have been too strong to allow any possible linguistic influence to emerge.[11] This line of reasoning would suggest that the participants' performance should be assessed only on the most suitable stimulus material, that is to say, only on those triads where the English and Italian results *taken together* show a fairly equally distributed split of choices between same-manner and same-path choices.[12] If linguistic relativity effects exist at all, it is in this kind of "perceptually best balanced" triads that such effects should show up most clearly. However, when the analysis was confined to the results of such triads, no significant difference between the two groups was found. A check made only on those 10 triads most relevant in this sense (i.e. triads # 3, 4, 6, 12, 13, 14, 17, 19, 22, 25) indicated that the difference was still not significant ($\chi^2 = 0.239$; $p = 0.6249$); another check made only on the five most relevant triads (i.e. triads # 3, 4, 13, 17, 22) came once again to the same conclusion ($\chi^2 = 0.042$; $p = 0.8383$).

A second possibility has to do with the fact that to perceive manner of motion only a few frames of the video-clip are sufficient, whereas with regard to a complete encoding of path of motion, more time is needed (Todd, 1983). As Gennari et al. (2002:71) rightly observe, especially when the similarity judgement task is run under *shadow condition*, as it was in the present study, where the participant's attention has to constantly switch between the auditory-motor task and the vision perception of the clips, it is possible that the perception of manner information is easier to encode quickly than that of path. This would result in participants from both groups being uniformly biased in choosing the same-manner variant, so that possible linguistic effects would not be able to show up. The experiment's results, however, are not particularly suggestive of this possibility, since, overall, the same-path variant was chosen more often than the same-manner variant across both linguistic groups.

A third possibility is that the participants could easily discover the common pattern of situations displayed by all triads (i.e. one variant clip always kept the manner of motion shown in the target unchanged, but not the path of motion; the opposite was true for the other variant) and quickly develop strategies for answering in a consistent way (e.g. always choose one and the same kind of variant, whichever of the two this was). In other words, the risk is that some responses may have been based more on conscious reasoning than on spontaneous, instinctive reaction to the stimulus. As already noted earlier in section 2.1.3, however, the simultaneous audio-motor + visual tasks must have considerably inhibited the possibility of conscious reflection on what was being viewed, since the participant was forced to perform two activities at the same time. One can at least say that the risk of a strategy-based responses has been lower than it would have been if the experiment had not involved the shadowing activity. Perhaps, a further precaution against such a kind of response could have been taken by placing filler items between the crucial triads of the test.

3. Conclusions

This study tested the Whorfian theory by comparing English and Italian speakers, both linguistically and non-linguistically, on the domain of manner of motion.

---

[11] Factors such as those just mentioned may well explain, for example, what was found in Experiment 1 of Hohenstein (2005), namely that for some of the triads involved in that experiment, the majority of the participants from both languages made the same kind of choice (Hohenstein, 2005:413–414).

[12] Because the numbers regarding the split of choices for every triad would be composed by English and Italian results taken together, they should not be influenced by possible linguistic relativity effects: any possible effect of this kind should be balanced out and eliminated.

With regard to the linguistic task, the data corroborated the previous findings of Cardini (2008) pointing to a significant difference between English and Italian speakers in the degree of linguistic salience for manner of motion. Once again, when compared to Italian speakers, English speakers produced a much larger number of manner verb tokens and types; the gap between the two groups was still substantial even when looking at the overall amount of elicited manner information, that is, manner information delivered not necessarily by the main verb of the phrase only, but also by some adverb, some gerundive form, or some PP. The verbal task also shed some light onto typological differences between the two language groups. By showing that Italian speakers provide manner of motion information predominantly through the main verb of the clause, it demonstrated the extreme flexibility with which Talmy's typological categorisation of S- and V-languages must be interpreted. At the same time, however, it also confirmed the overall kind of validity that that categorisation still seems to retain. The pattern of expression typically used by Italian speakers was still fundamentally different from that shown by English speakers. When the Italians expressed manner of motion, in nearly a third of the times they did so by adding some adverb, some gerundive form or some PP usually at the end of the sentence. In a classical satellite-framed fashion, by contrast, English speakers practically *never* delivered manner indication outside the main verb, unless they simply wanted to further elaborate the manner expressed by that verb. Moreover, while English speakers made an almost irrelevant use of path verbs, the amount of recorded path verbs among Italian speakers was greater than that of manner verbs.

With regard to the non-linguistic task, the similarity judgement task provided evidence against the Whorf hypothesis. The number of English preferences for same-manner choices was virtually identical to that for Italian; this is also true if one confines the analysis to that stimulus material where the two groups showed the widest gap in linguistic behaviour. It is important to point out that the present findings are consistent with those gathered by Gennari et al. (2002) in the shadow condition, that is, the only other linguistic relativity experiment on manner of motion which, in the present view, has actually checked a permanent kind of effects by suppressing possible inner speech during the time of the task. This allows us now to claim more confidently that *language does not seem to exert any steady bias on cognitive processes involved in the visual perception of motion events.*

As for the significance that the above findings may have in relation to the language-and-thought debate in general, they tend to support the thesis of those who view the language faculty as an independent modular system unable to affect any form of non-linguistic thought (Fodor, 1983). By contrast, they are less encouraging for those who view language and general cognition as integral to one another, a view that is suggestive of the existence in the speakers' mind of congruent correlations between the features of their particular language and those of cognitive mechanisms involved in their perception, attention, categorisation and memory. The findings do not necessarily deny validity to this second view of language, but they at least suggest that the degree of possible integration between language and cognition does have significant limitations.

Acknowledgements

References

Aske, Jon, 1989. Path predicates in English and Spanish: a closer look. Proceedings of the Berkeley Linguistic Society 15, 1–14.
Berlin, Brent, Kay, Paul, 1969. Basic Color Terms: Their Universality and Evolution. University of California Press, Berkeley.
Berman, Ruth A., Slobin, Dan I., 1994. Relating Events in Narrative: A Crosslinguistic Developmental Study. Lawrence Erlbaum Associates, Hillsdale, NJ.
Bohnemeyer, Jürgen, Eisenbeiss, Sonja, Narasimhan, Bhuvana, 2001. Event triads. In: Levinson, S.C., Enfield, N. (Eds.), 'Manual' for the Field Season 2001. Max Planck Institute for Psycholinguistics, Nijmengen, pp. 100–114.
Boroditsky, Lera, 2001. Does language shape thought? Mandarin and English speakers' conceptions of time. Cognitive Psychology 43, 1–22.
Boroditsky, Lera, Schmidt, Lauren A., Phillips, Webb, 2003. Sex, syntax, and semantics. In: Gentner, D., Goldin-Medow, S. (Eds.), Language in Mind: Advances in the Study of Language and Thought. MIT Press, Cambridge, MA, pp. 61–79.

Bower, Gordon H., Karlin, Martin B., Dueck, Alvin, 1975. Comprehension and Memory for Pictures. Memory & Cognition 3 (2), 216–220.

Cardini, Filippo-Enrico, 2008. Manner of motion saliency: an inquiry into Italian. Cognitive Linguistics 19 (4), 533–570.

Carmichael, Leonard, Hogan, H.P., Walter, A.A., 1932. An experimental study of the effect of language on the reproduction of visually perceived form. Journal of Experimental Psychology 15, 73–86.

Davidoff, Jules, Davies, Ian, Roberson, Debi, 1999. Colour categories in a stone-age tribe. Nature 398, 203–204.

Dukhovny, E., Kaushanskaya, M. Russian Verbs of Motion. Department of Psychology, University of California, Berkeley, Unpublished results.

Finkbeiner, Matthew, Nicol, Janet, Greth, Delia, Nakamura, Kumiko, 2002. The role of language in memory for actions. Journal of Psycholinguistic Research 31 (5), 447–457.

Fodor, Jerry A., 1975. The Language of Thought. Harvard University Press, Cambridge, MA.

Fodor, Jerry A., 1983. The Modularity of Mind. MIT Press, Cambridge, MA.

Gennari, Silvia P., Sloman, Steven A., Malt, Barbara C., Tecumseh, Fitch W., 2002. Motion events in language and cognition. Cognition 83, 49–79.

Gentner, Dedre, Goldin-Meadow, Susan, 2003. Wither Whorf. In: Gentner, D., Goldin-Meadow, S. (Eds.), Language in Mind: Advances in the Study of Language and Thought. MIT Press, Cambridge, MA, pp. 3–14.

Hohenstein, Jill M., 2005. Language-related motion event similarities in English- and Spanish-speaking children. Journal of Cognition and Development 6 (3), 403–425.

Hurlburt, Russell T., 1990. Sampling Normal and Schizophrenic Inner Experience. Plenum Publishers, New York.

Ibarretxe-Antuñano, Iraide, 2004. Language typologies in our language use: the case of Basque motion events in adult oral narratives. Cognitive Linguistics 15 (3), 317–349.

Imai, Mutsumi, Gentner, Dedre, 1997. A crosslinguistic study of early word meaning: universal ontology and linguistic influence. Cognition 62, 169–200.

Jackendoff, Ray, 1983. Semantics and Cognition. MIT Press, Cambridge, MA.

January, David, Kako, Edward, 2007. Re-evaluating evidence for linguistic relativity: Reply to Boroditsky (2001). Cognition 104, 417–426.

Jovanović, J., Kentfield, M. Manifold manner: an exploratory analysis of French and English verbs of motion. Department of Psychology, University of California, Berkeley, Unpublished results.

Levinson, Stephen C., 2003. Space in Language and Cognition Explorations in Cognitive Diversity. Cambridge University Press, Cambridge.

Levinson, Stephen C., Kita, Sotaro, Haun, Daniel B.M., Rasch, Bjorn H., 2002. Re-turning the tables: language affects spatial reasoning. Cognition 84, 155–188.

Li, Peggy, Gleitman, Lila R., 2002. Turning the tables: language and spatial reasoning. Cognition 83, 265–294.

Lucy, John A., 1992a. Grammatical Categories and Cognition: A Case Study of the Linguistic Relativity Hypothesis. Cambridge University Press, Cambridge.

Lucy, John A., 1992b. Language Diversity and Thought: A Reformation of the Linguistic Relativity Hypothesis. Cambridge University Press, Cambridge.

Lucy, John A., 1996. The scope of linguistic relativity. In: Gumperz, J.J., Levinson, S.C. (Eds.), Rethinking Linguistic Relativity. Cambridge University Press, Cambridge, pp. 37–69.

Lucy, John A., 2000. Introductory comments. In: Niemeier, S., Dirven, R. (Eds.), Evidence for Linguistic Relativity. John Benjamins Publishing Co., Amsterdam, pp. ix–xxi.

Naigles, Letitia R., Eisenberg, Ann R., Kako, Edward T., Highter, Melissa, McGraw, Nancy, 1998. Speaking of motion: verb use in English and Spanish. Language and Cognitive Processes 13, 521–549.

Özçalışkan, Şeyda, Slobin, Dan I., 1999. Learning how to search for the frog: expression of manner of motion in English, Spanish, and Turkish. In: Greenhill, A., Littlefield, H., Tano, C. (Eds.), Proceedings of the 23rd Annual Boston University Conference on Language Development, vol. 2. Cascadilla Press, Somerville, MA, pp. 541–552.

Papafragou, Anna, Massey, Christine, Gleitman, Lila R., 2002. Shake, rattle, 'n' roll: the representation of motion in language and cognition. Cognition 84, 189–219.

Papafragou, Anna, Li, Peggy, Choi, Youngon, Han, Chung-hye, 2007. Evidentiality in language and cognition. Cognition 103, 253–299.

Phillips, Webb, Boroditsky, Lera, 2003. Can quirks of grammar affect the way you think? Grammatical gender and object concepts. In: Proceedings of the 25th Annual Meeting of the Cognitive Science Society, Boston, MA.

Pourcel, Stèphanie, 2005. Relativism in the Linguistic Representation and Cognitive Conceptualisation of Motion Events Across Verb-framed and Satellite-framed Languages. University of Durham Press, Durham.

Pourcel, Stèphanie, 2009. Motion: a conceptual typology. In: Evans, V., Chilton, P. (Eds.), Language and Cognition in Space: The State of the Art and New Directions, Vol. 1. Equinox, London, Ch. 16.

Rosch Heider, Eleanor, 1972. Universals in color naming and memory. Cognitive Psychology 3, 337–354.

Slobin, Dan I., 1987. Thinking for speaking. Proceedings of the Thirteenth Annual Meeting of the Berkeley Linguistic Society 13, 435–444.

Slobin, Dan I., 1996. From 'thought and language' to 'thinking for speaking'. In: Gumperz, J.J., Levinson, S.C. (Eds.), Rethinking Linguistic Relativity. Cambridge University Press, Cambridge, pp. 70–96.

Slobin, Dan I., 2003. Language and thought online: cognitive consequences of linguistic relativity. In: Gentner, D., Goldin-Medow, S. (Eds.), Language in Mind. MIT Press, Cambridge, MA, pp. 157–191.

Slobin, Dan I., 2004. The many ways to search for a frog: linguistic typology and the expression of motion events. In: Strömqvist, S., Verhoeven, L. (Eds.), Relating Events in Narrative, vol. 2: Typological and Contextual Perspectives. Lawrence Erlbaum, Mahwah, NJ, pp. 219–257.

Slobin, Dan I., Hoiting, Nini, 1994. Reference to movement in spoken and signed languages: typological considerations. Proceedings of the Berkeley Linguistic Society 20, 487–505.

Talmy, Leonard, 1985. Lexicalization patterns: semantic structure in lexical forms. In: Shopen, T. (Ed.), Language Typology and Lexical Description, vol. 3: Grammatical Categories and the Lexicon. Cambridge University Press, Cambridge, pp. 36–149.

Talmy, Leonard, 1991. Path to realization: A typology of event conflation. Proceedings of the Berkeley Linguistic Society 17, 480–519.
Talmy, Leonard, 2000. Toward a Cognitive Semantics, vol. 2: Typology and Process in Concept Structuring. MIT Press, Cambridge, MA.
Todd, James T., 1983. Perception of gait. Journal of Experimental Psychology: Human Perception and Performance 1, 31–42.
Whorf, Benjamin L., 1956. Language, Thought, and Relativity: Selected Writings of Benjamin Lee Whorf. MIT Press, Cambridge, MA.

Born in Genoa (Italy) in 1963, **Filippo-Enrico Cardini** graduated at the University of Genoa in Modern Languages (English and German) in 1992. His past working experience has mainly involved jobs in the field of commerce and trade. Nevertheless, he has recently revived his interest for Languages and Linguistics by successfully completing an M.A. in Linguistics in 2003 (University of East Anglia), and, still in Linguistics, a Ph.D. in 2009 (Lancaster University).