Idea-based and Image-based Linguacultures

Evidence from American English and Mandarin Chinese

Durst-Andersen, Per; Barratt, Daniel

Document Version
Accepted author manuscript

Published in:
International Journal of Language and Culture

DOI:
10.1075/ijolc.17011.dur

Publication date:
2019

License
Unspecified

Citation for published version (APA):
Durst-Andersen, P., & Barratt, D. (2019). Idea-based and Image-based Linguacultures: Evidence from American English and Mandarin Chinese. *International Journal of Language and Culture*, *6*(2), 351–387. https://doi.org/10.1075/ijolc.17011.dur

Idea-based and Image-based Linguacultures: Evidence from American English and Mandarin Chinese

Per Durst-Andersen and Daniel Barratt

Journal article (Accepted manuscript*)

Please cite this article as:
Durst-Andersen, P., & Barratt, D. (2019). Idea-based and Image-based Linguacultures: Evidence from American English and Mandarin Chinese. International Journal of Language and Culture, 6(2), 351–387. https://doi.org/10.1075/ijolc.17011.dur

DOI: https://doi.org/10.1075/ijolc.17011.dur

Idea-based and image-based linguacultures:
Evidence from American English and Mandarin Chinese.

Per Durst-Andersen, CBS & Daniel Barratt, CBS

Abstract

In order to investigate whether or not cultural cognitive differences between Western and East Asian countries should be taken seriously we compared the empirical results from studies of perception and cognition involving primarily American and Chinese people to linguistic data from exactly the same areas in American English and Mandarin Chinese. What we found were systematic language parallels to the perceptual and cognitive differences found in empirical studies. Our linguistic analysis did not only reveal that the differences should be taken seriously, but also that it seems to be possible to trace them back to different perspectives involved: The Anglo-American culture has an idea-based perspective, while the Mandarin Chinese culture has an image-based perspective to what appears to be a common basis for both Americans and Chinese in all other respects. The difference in perspective is, for instance, reflected in the two very different writing systems.

1. *Introduction*

Empirical research from the cognitive sciences (e.g., Nisbett, Peng, Choi, & Norenzayan, 2001; Nisbett, 2003; Nisbett & Masuda 2003; Nisbett & Masuda 2006; Hedden et al., 2010; McKone et al, 2010) suggests that some significant cultural cognitive differences may exist between what in the literature is referred to as Westerners (people from Europe, North and South America, and Australia) and East Asians (people from China, Japan and Korea). The basic argument is that Westerners and East Asians have different cognitive styles. Westerners tend to cognize the world in a more *analytic* fashion: for example, they are more likely to think in terms of categories and rules, and to attend more to the focal objects in a visual scene. East Asians, on the other hand, tend to cognize the world in a more *holistic* fashion: that is, they are more likely to think in terms of contextual relationships and family resemblance, and to attend more to backgrounds and the relations *between* objects. In the existing literature one finds no single explanation for this distinction, but Nisbett (2003) refers to some historical and socioeconomic factors and adds some kind of evidence from the languages involved.

In this paper, we will examine the main results from experiments carried out by cognitive psychologists and cognitive neuroscientists which involve the way Westerners, primarily Americans, and East Asians, primarily Chinese, perceive and cognize things. After the examination of each area of perception and cognition (namely, object categorization and perceptual judgement; the perceptual process of visual scene analysis; and the cognitive processes of logical, causal, and dialectical reasoning), we will investigate the corresponding area in the mother tongues of American and Chinese people in order to see whether or not there are any parallels. Not only does it appear that all cultural cognitive differences found between people from the United States and China match differences between the American English language and the Mandarin Chinese languages, but it also appears that it is possible to trace the differences back to the same source: the Anglo-American culture seems to be idea-based, whereas the Mandarin Chinese culture seems to be image-based. This fundamental difference cannot be derived directly from the empirical investigations, but can be inferred from a detailed linguistic analysis of American English (hereafter English) and Mandarin Chinese (hereafter Chinese) which takes its point of departure in a new understanding of the word as being composed of

two pieces of content: *image* and *idea* mediated by an expression unit. In the concluding part of the paper, we will argue that (1) Americans and Chinese differ from one another in empirical tests, and (2) American English and Mandarin Chinese differ from one another in several respects due to the employment of different *perspectives* (cf. Verhagen 2007). We assume that Americans and Chinese share the same perceptual and cognitive basis as do all other people. Due to this, their mother tongues must be construed on more or less the same mental basis. But since the American linguaculture has an idea-based perspective and the Chinese linguaculture has an image-based perspective to what appears to be a common basis in all other respects, we obtain different findings in empirical tests and different findings in our linguistic analysis.

By linguaculture we understand how people in a certain speech community perceive things, how they think about things/cognize, how they act and interact, and how they communicate with one another. Considering the fact that people from cognitive psychology and cognitive neuroscience tend to focus on perception and cognition, the emphasis in our paper will be on these two areas. This cannot be otherwise since we take our starting point in the results from empirical experiments and not in the results from discourse analyses. We will, however, try to make up for this discrepancy when analysing American English and Mandarin Chinese. In the proceeding sections we shall review the research results on the areas mentioned above. The discussion of each area of perception and cognition will be followed by a systematic account of the corresponding area in language and ended by a summary which includes a comparison of the results from the cultural cognitive area and the results from the linguistic analyses. We chose to focus on English as an example of a Western language, and Chinese as an example of an East Asian language. It would have been preferable to include other languages as well, for instance, Japanese or Korean, because people from Japan and Korea are often used in the studies we refer to, but it was not possible to do that within the limits of one paper. However, our rationale was that United States and China are often viewed as prototypical members of two different cultural clusters: United States as a member of the Anglo cluster and China as a member of the so-called Confucian Asia together with, for instance, Japan and South Korea (cf. House et al., 2004; Chhokar et al., 2008).

When examining the empirical results from studies of perception and cognition we will use the notions that are used by the researchers themselves, but when describing the language parallels we will try to specify the notions in order to be able to describe the differences found in a more precise way.

2. *Basic differences between English and Chinese*

2.0 *Preliminary remarks*

In order to prepare the reader for the more detailed description of the English and Chinese languages and for the distinction between the idea-based and the image-based perspective, we shall start by shortly describing their writing systems, their grammatical systems and briefly mention the results of reading research. Although written language is subordinate to spoken language, both in function and in form, it seems to be relevant to include their writing systems because they differ fundamentally from one another. As we see it, it need not be a coincidence.

2.1 *The two writing systems*

The writing system of English and other Western languages is based on an *alphabetic principle*, i.e. words have content (e.g., *lift*), but are composed of letters which do not have any content at all (i.e., *l-i-f-t*). The function of letters is to specify the pronunciation of a word on the one hand, and to distinguish that word (*lift*) from other words (e.g., *left*, *loft*) on the other. Because of this distinctive function, words are arbitrary, conventional *signs* (cf. Saussure, 1916) (or symbols in Peirce's terminology, cf. Peirce, 1998). A complex idea can be expressed by a simple expression unit (e.g., *tax*) and a simple idea can be expressed by a complex expression unit (e.g., *university professor*).

The writing system of Chinese and other East Asian languages is based on a *logographic principle*. Each logograph or character represents either a simple idea or a percept. In this system, a character always corresponds to a lexical morpheme – i.e., an expression unit with a single piece of content – and always corresponds to a syllable in oral discourse. Characters are made up of strokes, varying from one to more than thirty strokes. The more strokes a character has, the more complex it is in content. Characters are not random arrays of strokes, but are constructed in a systematic way. Many

characters consist of what Peirce (1998) calls icons or indexes. Although only the basic characters are icons, being either pictograms or ideograms, it has to be stressed that in the majority of cases Chinese people assign a causal relationship between the meaning of the character and the various strokes of the character. Thus, characters can be regarded as *motivated signs* in the sense that they bear some resemblance to the corresponding content. Each word is traditionally composed of two syllables – i.e., two content units which are, in themselves, normally composed of several pictorial signs – together with some phonological information on the grounds that some characters are ambiguous (cf. Gao & Kao, 2002). Despite the fact that Chinese characters contain some information about pronunciation, Liu et al., 2002 showed that Chinese people rely more on the visual features than on phonological information. This finding goes against the traditional view that semantic activation is based on a phonological principle.

2.2 *The two grammatical systems*

In English, words are inflected according to their syntactic function in the sentence defined by the transitive or intransitive verb, e.g., *Can you give me some food?*, where *you* is subject, *some food* is direct object and *me* is indirect object, because *give* is a so-called ditransitive verb. The English verb has two aspectual forms (*is loving* and *loves*), four different tense forms (*gives, gave, has given* and *had given*), a sharp distinction between definite and indefinite articles (*the man* and *a* man) and a diathesis distinction between active and passive (*he gave the book* vs. *The book was given to him*). English has *that*-sentences (*He said that he would come*) and relative clauses (*He gave the book which was written by George Orwell*).

In Chinese, one will not find inflection at all. One will not find definite or indefinite articles, case forms, or tense forms (see Lin 2006). Moreover, one will not find *that*-sentences or relative clauses, because there is not a character for *that* or *which/who*. We find a Chinese construction that functions as a relative clause in English, but the construction is not a clause and there is no pronoun referring to a preceding noun. The only distinction found in Chinese which resembles something in so-called Western languages is the distinction between the perfective and the imperfective aspect (cf. Li & Thompson, 1981; Wu, 2005; Xiao & McEnery, 2004; Lin, 2003). If a sentence denotes a state or

an activity it involves a single verb as we know it from English, French or German, but if a sentence denotes an action (e.g., "Give smb. food" or "Go to Paris") it is seldom composed of one verb, because Chinese is a language with so-called *serial-verb constructions*. *Can you give me some food?* could correspond to the following Chinese utterance: "you go take bring come give me food"? Chinese names every activity involved almost as a photographic representation of each movement: the hearer is asked by the speaker to go away from him, to take some food, to bring it back, to approach the hearer so that the food goes from the speaker to the hearer (see Peyraube, 2006).

If we look at communication patterns being practised in the United States and in China, they also seem to differ in a fundamental way. Hall (1976/1997) and Ting-Toomey (1999) distinguish between low-context communication (self-face concern, direct style, speaker-oriented style and verbal-based understanding) and high-context communication (mutual-face concern, indirect style, listener-oriented style and context-based understanding) and place the United States and China within the frameworks of low-context and high-context communication respectively.

2.3 *Reading research*

Research suggests that while reading Western and Chinese people scan a sentence in the same way – i.e., from left to right and from top to bottom (see Castelhano & Rayner, 2008; Tsiang & Chen, 2008), but it is important to note that the same procedure is applied to completely different things. Western people scan letters comprising words, whereas Chinese people scan small pictures. Furthermore, each picture/character is read from top to bottom. We turn the reader's attention to the fact that English uses vertical strokes (i.e., Roman numerals) for the numbers 1, 2, and 3, i.e., going from left to right, whereas Chinese use horizontal strokes going from top to bottom. Words yield a cluster, but not a picture which in a certain way resembles its content. As suggested above, a writing system is always derived from its oral system one way or the other. In Western languages, the letters making up the alphabet specify different pronunciations, i.e., is based on the sense of hearing. Although this is not the case in Chinese, it is reasonable to assume that the Chinese writing system reflects something. One possibility is that Chinese characters are based on the sense of vision and thus reflect a perceptual basis of Chinese grammar. If Chinese grammar is based on images – in contrast to the ideational basis

of Western languages – it seems plausible that the Chinese way of reading and writing characters reflects the way they look at a picture and vice versa.

3. *Categorizing from the point of view of an idea or an image*

3.1 *Making sense of objects*

When considering the variety of cognitive processes potentially susceptible to cultural influences, a good place to start is where the child starts; that is, the capacity to make sense of the objects in their environment. The working hypothesis is that Westerners tend to make sense of objects in terms of the assignment of *categories* and the application of *rules*, whereas East Asians tend to make sense of objects by looking for *relationships* and *similarities*. Two studies are commonly cited in the literature.

In the first study (Chiu, 1972; also see Unsworth, Sears, & Pexman, 2005), American and Chinese children were shown pictures of three target objects – for example, a cow, a chicken, and some grass – and asked, 'Which two target objects belong together?' (see Fig. 1a). The American children tended to adopt a more analytical cognitive style and grouped the objects in terms of *categories*: for example, they were more likely to link the cow with the chicken (both examples of basic level categories) on the grounds that both the cow and chicken could be classified under the (superordinate) category of 'animal'. The Chinese children, on the other hand, tended to adopt a 'relational-contextual' cognitive style and grouped the objects in terms of *relationships*: for example, they were more likely to link the cow with the grass on the grounds that cows typically stand on and eat grass.

In the second study (Norenzayan, Smith, Kim, & Nisbett, 2002, Study 2), European American, Asian American, and East Asian (Chinese and Korean) students were shown pictures of two groups of objects and one target object, and asked, 'Which group of objects does the target object belong to?' (see Fig. 1b). The Western participants tended to think in terms of *rules*: they were more likely to link the target flower to the four flowers in group 2 on the grounds that there was a single common denominator (namely, a straight stem). The East Asian participants, on the other hand, tended to think in terms of *similarities* (or family resemblance): they were more likely to link the target flower to the four flowers in group 1 on the grounds that there were a greater number of shared perceptual features.

(In group 1, three out of the four flowers have oval-shaped petals, a single centre, and/or an offshoot with leaf; whereas in group 2, each of these features occurs only once.)

[Figure 1]

3.2 *Language parallels – The word as an image-idea pair*

It is impossible to spot essential differences in categorization and naming between languages, if one takes the starting point in Saussure's definition of the linguistic symbol as two-sided, i.e. as consisting of an expression unit and a content unit in which there is an arbitrary relationship between the two sides (Saussure, 1916). If, however, one adopts Peirce's trichotomic view of a symbol in which any expression unit, called *representamen*, has two sides of content, called *object* and *interpretant* (cf. Peirce, 1998), it becomes possible to find differences and to consider their motivations. Peirce was a phenomenologist in his point of departure and reality as such was not described, for instance, from the point of view of a positivist. He called reality zero-ness indicating that he was not interested in reality as such. The *object* is firstness, the *interpretant* secondness and the *representamen* which mediates the object and the interpretant is thirdness from the point of view of Peirce. Since Peirce's category of firstness has to do with experiences, his *object* is more or less similar to our *image* covering all experiences connected to the human senses, i.e. sight, hearing, touch, taste and smell. Since Peirce's category of secondness has to do with the effect of the experience in the human mind, his *interpretant* is directly related to our *idea* involving the mental understanding of the experiences of the human body. Since Peirce's category of thirdness mediates secondness and firstness by viewing them in totality from an external point of view, his *representamen* is similar to our unit of expression which is argued to mediate the idea side and the image side, i.e. two different kinds of content (Author, 2009; 2011: 143-148).

The double nature of the content side of linguistic symbols explains why and how people can relate entities in reality and in their mind to linguistic entities, and vice versa. When receiving a visual picture of a certain entity, e.g., "book", in a normal world anyone is capable of describing it as <an object consisting of bound pages with words written on them or pictures drawn on them> and

applying the right name, e.g., *book* – the linguistic equivalent to the so-called *transduction problem* (Barsalou, 1999). Similarly, hearing a description of an entity any person will be capable of drawing a picture of it and at the same time naming it – the linguistic equivalent to the so-called *grounding problem* (Harnad, 1990). And last, but not least, hearing a word any person is capable of drawing a picture that matches its referent and giving a description of it. It is not possible to account for all this by using the traditional approach which involves an expression unit and a content unit.

The Peircean approach also makes it clear that the arbitrary relationship is found between the expression unit and its image side, but not between the expression unit and its idea side. Note that Peirce's icon, index and symbol match three different types of semiotic relations between the expression unit and its object – in our terminology image. In other words, from the point of view of English it seems completely arbitrary that the image of a "golf course" is called *golf course*, but it seems to be motivated that the idea behind a "golf course" <a location for rule-guided golf activity> is reflected in the name *golf course*. In the same way, from the image point of view it is completely arbitrary that the situation "somebody did something with the effect that he is invisible from other persons on purpose" is called *he is hiding somewhere*. In Chinese the same is called *duǒ-jìn* which means <go away-enter>. Here the idea of being invisible for others is not expressed, but rather the image side of "hiding" seen from the agent's own perspective: first, the person goes away, then he enters some place, and the effect of this is that he is invisible to the observer, but this is not named in Chinese. We may certainly argue that in the Chinese case the relation between the single verb and its image content is arbitrary, but when they are put together in that particular order it becomes motivated from the point of view of their respective image content.

3.3 *Idea-based vs. image-based*

The Western tendency to think in terms of categories and rules, and the East Asian tendency to think in terms of contextual relationships and family resemblance seem to be reflected in language. English shows a clear tendency to be *idea-based* in their naming strategies, whereas Chinese shows a clear tendency to be *image-based*.

Most nouns can be said to have both an image side – a prototypical pictorial representation – and an idea side – a description of the object's category and function. Idea and image go together: one can give a description of a chair and make a drawing of it and tell how it feels to touch it and sit on/in it. When a language ascribes a name to a new chair, however, it has to make a fundamental choice between the idea and the image. Languages with a preference for the idea-content tend to have a single collective concept for classifying a variety of different chairs and a variety of different bowls: for example, different chairs are united in terms of serving a common function (providing a seat for one person, usually with a back and four legs) and different bowls are united in terms of serving a common function (providing a container for various sorts of stuff). In contrast, languages having a preference for the image-content lack a collective name and tend to have different words for different chairs and for different bowls: for example, chairs in a dining room look differently from chairs in a living room which look differently from chairs in a café which again look differently from chairs in an office; and, similarly, a bowl with salad looks differently from a bowl with sugar that looks differently from a bowl with punch and so forth. This is, for instance, the case in French where artefacts, but not nature facts, are named by the hyponyms: *chaise* 'dining chair' and *fauteuil* 'easy chair' (and these are used to derive further subtypes of chairs that look like "dining chairs" and "easy chairs"); and, for instance, *saladier* 'salad bowl', *sucrier* 'sugar bowl', and *bol* 'punch bowl' (and these are used to derive further subtypes of bowls that look like "salad bowls", "sugar bowls" and "punch bowls"). In French it is extremely difficult to find a name for the collective concept of an artefact. The logic seems to be that French have different names for objects that look differently. Since "chair" as such or "bowl" as such are not visible, only manifestations of these two concepts, i.e. "dining chairs"/"easy chairs" and "salad bowls"/"sugar bowls"/"punch bowls", are named (cf. Baron & Herslund 2005; Herslund & Baron 2003).

The image-based categorization in Chinese is manifested in a number of ways and in a different way to French. Chinese, for instance, has a name for "chair", but Chinese people prefer to use the hyponyms in situations where Americans would use the hypernym *chair*. Since the hyponyms have a character that denotes an image ("rocking", "bar", "wooden", etc.) placed in front of the character that denotes the idea of chair, Chinese shows that it is image-based. At a general level,

Chinese has ‘genuine’ nouns for collective concepts relating to natural objects: for example, *niú* ‘cattle’, *yáng* ‘sheep’, *niǎo* ‘bird’, and *rén* ‘person’. At a more specific level, however, Chinese has nouns composed of the corresponding hyponyms, as appears from (1):

(1) *mǔ-niú* ‘cow’ vs. *gong-niù* ‘bull’; *mǔ-yáng* ‘ewe’ vs. *gong-yáng* ‘ram’; *mǔ-jī* ‘hen’ vs. *gong-jī* ‘cock’; and *nǚ-zǐ* ‘woman’ vs. nán-zǐ ‘man’.

In the case of four-legged animals, females are denoted by the hyponym *mǔ* and males are denoted by the hyponym *gōng*. In the case of two-legged animals, females and males are denoted by the hyponyms *cí* and *xióng* respectively, whereas for (two-legged) humans the corresponding hyponyms are *nǚ* and *nán*. This dependence on hyponyms extends to other types of nouns as well. Whereas English has nouns for collective concepts such as “parents” and “grandparents”, Chinese has composite terms, cf. (2)

(2) *fù-mǔ* (lit. ‘father-mother’); *zǔ- fù-mǔ* (‘ancestor-father-mother’).

In summary, when it is impossible to see a figure against a ground, one will not find a single noun but instead a collection of the constituent hyponyms. When it is possible to see a figure against a ground, one will find a noun but at the level of the constituent hyponyms one will find distinctions that seem to go back to differences in perception (at the image level) rather than differences in cognition (at the idea level): for example, female animals with four legs will be differentiated from female animals with two legs, rather than being grouped together under a general category for female which is understandable from the point of view of the idea but not directly observable (For experimental evidence for the distinction between perceptually-based and conceptually-based recognition, see Gerlach, 2008; 2009).

As already pointed out, Chinese does not know of an article system a la the English one between *a book* and *the book* which distinguishes pieces of new information (signalled by *a*) from pieces of old information (signalled by *the*) to the hearer (word order is the only way Chinese can distinguish new from old information). Instead of an article system, Chinese has a system of around eighty classifiers, which in case of reference to the real world (as opposed to an imagined world) have to occur before a given noun as, for example the following five classifiers, in (3):

(3) *zhāng* (used for flat things such as pictures), *běn* (used for objects with many pages such as books), *jiā* (used for container-like objects such as restaurants), *zhī* (used for long, thin inflexible objects such as sticks), *gēn* (used for not thin objects such as flag poles), and the more general *ge* (used in connection with pure quantification, e.g., *wǒ kàn le yī ge diànyǐng*, 'I have seen a movie').

The important question is why there are almost eighty different classifiers in total. If one looks closely into their semantics, it seems to be evident that their content is based on looking at specific pictures of objects which then have been transformed into images or percepts; that is, into a form that can capture all possible pictorial manifestations (for a similar view on classifiers in general, see Aikhenvald, 2000; Allan, 1977: 308f). Thus, once again, we observe a specific Chinese interest in image-based thinking, not in idea-based thinking. The categorization is made on the basis of images, not on the basis of ideas.

In English one finds an article system which distinguishes new information from old information by using the indefinite article and definite article, respectively. Pieces of new or old information cannot be visualized, but can only be understood: new information reflects an established non-identity between the speaker's and the hearer's mental universe, while old information reflects an established identity between them.

3.4 *Summary*

In short, what we observe in the Chinese language parallels what empirical findings have shown. Chinese children place "cow" and "grass" in the same category, because they appear in the same picture. East Asian participants place the target flower among other flowers that in general look in the same way. They seem to apply an image-based approach. Contrary to this, English seems to be idea-based in its naming and categorization which also parallels what we just saw in the investigations just examined above. American children place "cow" and "hen" in the same category, because they share the feature 'animal', and American participants place the target flower among flowers that all without exception share the same feature 'straight stem', but look very different. They thus seem to have an idea-based approach. This means that the cognitive differences we find in empirical research among

Western and East Asian participants have exact parallels in the participants' mother tongues. This tells us that the notion of image has another status in the Chinese society than in the English one. This may explain why Chinese menus involve images of the food that can be ordered at a restaurant and why Chinese advertisements involve far more visual elements, including elements from fantasy world, than American advertisements do.

4. *Perceiving an entity in an absolute or relative way*

4.1 *Perceptual judgements*

The second psychological faculty of interest is the capacity to make perceptual judgements; that is, judgements about a basic feature of a focal object such as orientation or size. In this case, the working hypothesis is that East Asians are more *field dependent* than Westerners. The notion of field dependence can be construed both positively and negatively: Positively, in the sense that East Asians are more able to bind the focal object (the figure – a notion that goes back to the famous Danish psychologist Edgar Rubin (1915) that was revived by cognitive linguists in the eighties (cf. Lakoff 1987)) with its surrounding field (its ground) in a holistic fashion; and negatively, in the sense that East Asians are less able to separate the focal object (figure) from its surrounding field (ground) in an analytic fashion. Two types of perceptual judgement task are commonly cited in the literature.

In the Rod-and-Frame Test (RFT; Witkin, 1967; Ji, Peng, & Nisbett, 2000), the task is to judge the verticality of a rod and to ignore the orientation of a surrounding frame (see Fig. 2c). The main finding can be described in terms of field dependence: in comparison with Western participants, East Asian participants are less accurate at the task because they are less able to separate the focal object (the rod) from its surrounding field (the frame). Similarly, in the Framed-Line Test (FLT; Kitayama, Duffy, Kawamura, & Larsen, 2003; Hedden et al., 2008), the stimulus is a picture of a square frame intersected by a vertical line (see Fig. 1d). In this case, the main finding is that Western participants are more accurate at the absolute task (drawing a new line that is identical in length to the original line), while East Asian participants are more accurate at the relative task (drawing a new line that is the correct length in proportion to the surrounding frame).

Another example of a perceptual judgement task uses Navon figures (i.e., a large letter composed of smaller letters). When presented with Navon figures, Western participants tend to recognize the small letters and show more local (analytic) patterns of attention, whereas East Asian participants tend to recognize the large letters and show more global (holistic) patterns of attention (e.g., McKone et al., 2010).

The notion of field dependence is also supported by two of the studies on scene perception involving verbal description and eye tracking (Masuda & Nisbett, 2001; Chua, Boland, & Nisbett, 2005; described below). Both studies contained a follow-up object-recognition phase in which the participants were reshown the focal objects against either the same or a new background. The East Asian participants were less likely than the American participants to recognise the focal objects when presented against new backgrounds, suggesting that East Asians tend to 'bind' objects with backgrounds in both attention and memory.

4.2 *Language parallels*

The English and Chinese languages differ fundamentally from one another in the way they describe various domains including those which are reflected in the Rod-and-Frame Test and the specific task involved in the modified FLT-version found in Hedden et al. (2010). While English uses verbs such as *have* and *be*, Chinese has a specific device which binds figure and ground, but differs with respect to starting point. Cognitive linguists such as Langacker, Lakoff and Talmy use these notions, but they have never been applied to *have*- or *be*-verbs in, e.g., English, Russian or Chinese.

Traditionally, linguists distinguish between 'Be-languages' and 'Have-languages' according to the way a language expresses "possession": for example, an English speaker would say *I have a computer*, whereas a Russian speaker would say *Computer exists with me*. The distinction between English as a Have-language and Russian as a Be-language is crucial, because it also affects what is normally called "experience". English uses either *have* as in *I have a need for smth* or *I have a desire for smth* or *be* as in *I am fine* or *I am cold* etc., whereas Russian consistently uses another Be-verb plus the dative case for expressing experience: *For me smth is necessary/desirable* or *For me it is*

cold. The crucial thing is that English systematically uses a human being as subject, whereas Russian systematically uses the given figure as subject.

In Chinese one will not find a verb as *have* or *be*, but instead another type of distinction which is equally important and is reminiscent of Russian. The distinction is purely perceptually based. With respect to external (observable) reality Chinese has two so-called *perspectivizers* which are used to describe physical locations (see Author 2011: 41-43). The verb *zài* is used to give a figure-ground perspective, where the notions of 'figure' and 'ground' correspond roughly to the notions of 'focal object' and 'background context', respectively (4):

(4) a. *tā zài dàxué.* 'He (figure) is at the university (ground).'

b. *tā zài dàxué gōngzuò.* 'He (figure) is working at the university (ground).'

Conversely, the verb *yǒu* is used to give a ground-figure perspective (5) (cf. Hu & Pan, 2008):

(5) a. *wǒmen yǒu shíjiān.* 'We (ground) have time (figure)'.

b. *wǒmen méi yǒu háizi.* 'We (ground) don't have children (figure)'.

c. *wǒ jiā yǒu liù kǒu rén.* 'My family (ground) consist of six people (figure)'.

d. *yǒu rén zhǎo nǐ* 'There (ground) is a person (figure) who is looking for you'.

We notice that Chinese does not distinguish between animate and inanimate entities, i.e. people and places are treated identically as potential figures and grounds. This means that *yǒu* can be used to describe both location and possession, thus showing that the Chinese distinction is different from that. It is connected to two different perspectives that are not found in American English. Note that both perspectivizers may have animate and inanimate entities as figure and ground.

The use of perspectivizers is not restricted to state descriptions involving location and possession. We find the same phenomenon in action descriptions, where Chinese uses a verb instead of a preposition in order to give a certain perspective. A good example of this would be *gěi* which is the verb for "X give smth. to smb." Besides its normal use as a verb as in (5), it is also used for the English preposition *to* as in (6), where the position of *gěi* before the actual main verb *ná* 'take' indicates a certain perspective, namely that a certain figure (some ice cream) is placed on a ground, i.e. the little boy.

(5) *shū shū zài qù ná yì diǎnr gěi nǐ.*

uncle again go take a little bit give you

'Uncle can give you more ice cream.'

(6) *shū shū zài gěi nǐ ná yì diǎnr.*

Uncle again to you take a little bit

'Uncle can take more ice cream to you.'

The introvertive perspective given by *gěi* in (6) has its counterpart in the particle *de* which expresses the extrovertive perspective in which a certain figure has been removed from its ground. Two other verbs share this ability to appear as perspectivizers although they are born as full verbs. It concerns, for instance, *qù* 'go' and *lái* 'come' that as perspectivizers mark the speaker's starting point or end point, respectively (Zhang 2007). With respect to internal reality, Chinese uses neither *zài*, nor *yǒu* – they are only used for descriptions of situations in external reality. Experience is conveyed by zero (cf. 7a), while categorization is expressed by *shì* (cf. 7b):

(7) a. *tā Ø hǎo le.* 'He is fine (now).'

b. *tā shì kuàijìshī.* 'She is an accountant.'

The important thing is, however, that Chinese people have to relate figure to ground or ground to figure when they describe external reality. This is surely not the case with American people.

4.3 *Summary*

The fact that it is easier for American students to solve absolute tasks, while it is easier for Chinese students to solve relative tasks find a natural explanation in their respective mother tongues. American children are not forced by English to compare either figure-ground or ground-figure. English does not know of figure-ground relationships. In Chinese there is no way to avoid the two varieties of this relative viewpoint when speaking of external reality. This means that in Chinese it is impossible to leave out the context and focus entirely on the central item (what is often referred to as minimalism within art and design). A perspective will always include two things, in this case the notions of figure and ground. This is not tantamount to saying that language determines thought, but the incorporation of the English and Chinese languages accentuates the relevance of the distinction between so-called analytical or abstract way of thinking in Western countries and the so-called holistic way of thinking

in East Asian countries. We would suggest to name the two opposites idea-based thinking and image-based thinking, respectively: an idea will always imply focus on one entity, be it abstract or concrete, whereas an image will always imply a frame consisting of a figure and a ground. By using the notions of idea-based and image-based it is possible to incorporate cognition and language at the same time.

5. *Perceiving a complex scene in a local or global way*

5.1 *Introducing perception of visual scenes*

Many of the cultural differences in what is referred to as analytic and holistic ways of thinking seem to boil down to a Western tendency to attend more to objects, figures, and an East Asian tendency to attend more to contexts, grounds. A follow-up question then is: Are these attentional tendencies reflected in the way that Westerners and East Asians actually look at the world; that is, in the concrete operations of the human visual system? The working hypothesis is that Westerners attend more to focal objects in a visual scene, whereas East Asians attend more to the relationships between objects and to background elements. Experimentally, this hypothesis has been investigated using a variety of stimulus materials including: visual scenes, presented in the form of either still images (e.g., Chua, Boland, & Nisbett, 2005) or dynamic images (e.g., Masuda & Nisbett, 2001); facial expressions (e.g., Masuda, Ellsworth, et al., 2008); and works of art (e.g., Masuda, Gonzalez, Kwan, & Nisbett, 2008). The corresponding experimental paradigms can be organised in chronological order and by increasing directness of approach.

5.1.1 *Verbal description*

A first and relatively indirect method for investigating the perception of visual scenes is verbal description (either during or after observation of those scenes). In a study by Masuda and Nisbett (2001), for example, European American and Japanese participants were presented with short animated vignettes of underwater scenes. Each vignette contained both salient focal objects (e.g., a large fish) and various background elements (e.g., smaller fish, snails, plants, and bedrocks). Afterwards, the participants were asked to verbally describe what they had seen, and the resulting descriptions were transcribed and coded. The main finding was that the American participants

reported more information about focal objects (and were more likely to mention the focal objects first), whereas the Japanese participants reported more information about the relationships between objects and about the background elements.

In a more recent study, Senzaki, Masuda, and Ishii (2014) have combined the verbal description method with the eye tracking method described below. In this study, European Canadian and Japanese participants were presented with modified versions of the animated vignettes. When the task was observation only, there were no significant differences in the eye movement patterns between the two cultural groups. When, however, the task was to verbally describe the vignettes (by constructing narratives), both the eye movement patterns and the verbal descriptions differed in the manner expected, with the Canadian participants paying more attention to focal object information and the Japanese participants paying more attention to contextual information.

5.1.2 *Change detection*

A second method for investigating the perception of visual scenes can be broadly defined as change detection. In the flicker paradigm developed by Rensink, O'Regan, and Clark (1997), the task is to detect the change – or 'spot the difference' – between a pair of alternating images as quickly as possible. (The failure to detect such changes relates to the phenomenon of change blindness; Simons, 2000.) In a cultural version of the flicker paradigm, Masuda and Nisbett (2006; Study 1) showed European American and East Asian (Chinese, Japanese, and Korean) participants pairs of still images of realistic industrial scenes. Here the main finding was that the American participants were faster at detecting changes in focal object information, whereas the East Asian participants were faster at detecting changes in contextual information. In the interests of ecological validity, Masuda and Nisbett (2006; Studies 2-3) also tested dynamic scenes and obtained equivalent results.

More recently, Masuda, Ishii, and Kimura (2016) have combined the change detection/flicker paradigm with eye tracking. This time, European Canadian and Japanese participants were presented with a number of sporting scenes, some of which (unbeknownst to the participants) did not contain a change. In the change trials there were no significant differences in the eye movement patterns

between the two cultural groups, whereas in the no-change trials the eye movement patterns of the Canadian and Japanese participants differed accordingly.

5.1.3 *Eye tracking*

The most direct and 'online' approach to investigating the visual system is provided by eye tracking, a method for recording eye movements in terms of fixations and saccades. Following the premotor theory of visual attention, eye movements can be regarded as indexes of visual attention to the extent that eye movements reliably follow shifts of visual attention (see Rizzolatti, Riggio, Dascola, & Umiltá, 1987). In a founding study, Chua, Boland, and Nisbett (2005) presented both European American and Chinese participants with manipulated photographs of a focal object (either an animal or a non-living entity) against a relatively complex background. The main finding was that the American participants fixated the focal objects sooner (on average, by 118 milliseconds) and for longer durations. In contrast, the Chinese participants made more fixations on (and saccades to) the backgrounds. Interestingly, no culture differences were observed for the first 300 to 400 milliseconds of stimulus presentation.

Despite the conclusions of this study, it is not yet clear whether cultural differences in eye movements truly exist. In two subsequent eye tracking studies, Rayner et al. (2007) only found evidence of cultural differences for a subset of their stimuli, while Evans et al. (2009) failed to replicate the results of the Chua et al. study using the same stimulus set. (For a debate between the two competing groups of researchers, see Boland, Chua, & Nisbett, 2008; Li, Williams, Cave, Well, & Rayner, 2008. The recent studies by Senzaki, Masuda, & Ishii, 2014, and Masuda, Ishii, & Kimura, 2016, described above suggest that cultural differences in eye movements only emerge in certain conditions.)

A similar pattern of results can be found for the perception of facial expressions. When judging facial expressions of emotion, for example, East Asians are more likely to refer to, and look at, the social context in which those expressions occur, possibly reflecting a greater concern for group emotions (e.g., Masuda, Ellsworth, et al., 2008; Ito, Masuda, & Hioki, 2012). It remains to be seen whether cultural differences in global and local attention extend to face processing itself, with

Westerners and East Asians applying different patterns of attention to the eye, nose, and mouth regions.

5.1.4 *Aesthetics: Preferred way of framing things*

A fourth, and somewhat miscellaneous, method tackles the issue from the opposite direction. Instead of showing participants visual scenes and measuring their perception in some way, one can investigate how participants (or artists) actually create visual scenes themselves, thereby using aesthetic preferences as an indirect measure of which elements are perceptually important. For example, Masuda, Gonzalez, Kwan, and Nisbett (2008; Study 2) conducted a two-part study with American students representing the West, and Chinese, Japanese, Korean, and Taiwanese students representing the East. In the picture-drawing task, the participants were asked to draw a picture of a landscape which included at least five elements (a house, a tree, a river, a person, and a horizon), but were free to draw additional elements if they wished. In the photograph-taking task, the participants were given a camera and asked to take portraits of a model. In comparison with the American participants, the East Asian participants drew the horizon line higher in the picture plane and included a greater number of objects; they composed portrait photographs in which the model covered a proportionately smaller area and the background covered a proportionately larger area. In a complementary study, Masuda and colleagues (2008; Study 1) examined a sample of paintings from the history of Western and East Asian art and found an equivalent pattern of results. We would like to turn the reader's attention to two corroboratory studies on culture and framing preferences (Nand, Masuda, Senzaki, & Ishii, 2014; Senzaki, Masuda, & Nand, 2014.), which show that the two different ways of framing things can be seen in art history as well as in the early development of children.

5.2 *Language parallels*

The above research suggests that Westerners attend more to foreground (focal) objects in a visual scene, whereas East Asians attend more to backgrounds and the relations *between* objects. Potential parallels in Western and East Asian languages can be found by examining *word frequency* (roughly analogous to how frequently and how long participants look at certain regions of a visual scene) and

word order (roughly analogous to the temporal sequence in which participants look at those regions). Additional evidence can be found by investigating how the speakers of the respective languages perform activities such as introducing themselves, writing addresses, and presenting time.

5.2.1 *Word frequency*

Developmental studies have observed the ways that Western and East Asian mothers interact with their children, with a particular focus on the use of language (for a study of American English- and Mandarin Chinese-speaking families, see Tardif, Gelman, & Xu, 1999; for a study of American and Japanese families, see Fernald & Morikawa, 1993). One of the main findings is that Western mothers tend to direct their child's attention towards objects (for example, the child's toys) and their properties. In language, this tendency is revealed in the more frequent use of nouns: for example, 'Look at the *car*. The car is *red*.' East Asian mothers, on the other hand, tend to direct their child's attention towards the field and the relationships *between* objects. In language, this tendency is revealed in the more frequent use of verbs. A noun normally names a focal object (figure), whereas a verb normally names the relationship between objects and the background context (a figure-ground relationship).

5.2.2 *Word order in text-initial utterances*

If we take a careful look at the very beginning of an English/Danish narrative and contrast it with the very beginning of a Chinese/Russian narrative, there are crucial differences. One could ask: Why look at the very beginning of a narrative? The answer is simple: The very first utterance provides us with the best conditions for making a comparison between languages – the speaker starts from scratch without being under the influence of a preceding utterance that would necessarily dictate a certain word order. Let us suppose that Figure 2 marks the beginning of a story that has to be told in English, Danish, Chinese, and Russian (Danish is included in order to show the three logical possibilities – it could also have been Norwegian or Swedish).

[Figure 2]

Being exposed to the picture in Figure 3 and being given the task of telling a story it is demonstrated by Durst-Andersen(2011, 2012) that English and Danish speakers begin their story in a different way from that of Chinese and Russian speakers (cf. 8, 9, 10 and 11):

(8) *In 1945 in Moscow five people were sitting around a table in a small café.*

(9) *Der sad fem personer omkring et bord i en lille café i Moskva 1945.*
There sat five persons around a table in a small cafe in Moscow 1945

(10) *zài yī jiǔ sì wǔ nián mòsīke yī jiā xiāo kāfēi guǎn lǐ,*
be 1945 year Moscow one CL small café house inside,
wǔ ge rén wéi zuò zài yī zhāng zhuōzi pang
five CL people round sitting CL table besides

(11) *V 1945-m godu v Moskve v malen'kom kafe za stolom sidelo pjat' čelovek.*
In 1945 year in Moscow in small café around table sat five people

When putting all pieces of information – corresponding to figure, ground and scene – into a single utterance, American English speakers begin the utterance by describing the scene, *In 1945 in Moscow*, continue by describing the figure, *Five people were sitting*, and finish by describing the ground, *around a table in a small café*, cf. (8). If we stick to what has been examined in the previous sections, we conclude that English speakers proceed 'from-figure-to-ground'. Scene is irrelevant when speaking of a picture, since it is normally not visible in a single picture. Danish people would say, cf. (9): There were sitting five people; around a table in a small café; in Moscow in 1945, or in short: Figure – Ground – Scene. Again we observe the order from figure to ground which reveals that figure is the starting point. In contrast, cf. (10), Chinese speakers begin the utterance by describing the scene, *In 1945 in Moscow*, but continue by describing the ground, *in a small café*, and end by describing the figure, *five people were round-sitting besides table*. In other words, if we stick to the relative ordering of figure and ground, Chinese speakers proceed in the opposite direction 'from-ground-to-figure'. Russian people would say, cf. (11): In 1945 in Moscow; in a small café around a table; five people were sitting, or in short: Scene – Ground – Figure. This is identical to the Chinese way. Note that the three types of utterances represent the normal, unmarked *text-initial word order* in American English,

Danish, Chinese and Russian. This does not of course exclude the possibility that in some cases English speaking persons would prefer to split the entire utterance into two utterances.

The crucial point is that in the case of text-initial word order one will not find the American English word order in Chinese or Russian, and, similarly, one will not be able to find the Chinese and Russian word order in American English or Danish. Although we observe three different framing strategies corresponding to the American English, Danish and Chinese/Russian ones, it is crucial to note that if we isolate the order of figure and ground, then there are only two ways: Either one starts with the ground as in Chinese and Russian or with the figure as in American English and Danish. Note that if we were to continue the story by saying "They were sitting in the corner", the word order of figure and ground would be exactly the same in the four languages.

The important question is: What do different word order patterns reflect at this specific place? If text-initial word order should reflect something – and we think it does besides it represents a certain framing strategy – it must necessarily be the way in which the speaker looks at the total picture being described and/or the way in which the hearer is supposed to look at it. In short, it must reflect the speaker's and/or the hearer's perception strategy.

5.2.3 *Other pieces of evidence for different perception strategies*

When looking for other possible areas which may reveal additional pieces of evidence for applying different framing strategies (and hence also different perception strategies), three things come to our minds. The first one is how you introduce yourself. The second one is how to write a complete address on an envelope. And the third one is the exact ordering of various time specifications. Let us consider them in turn.

Introducing yourself in American English and Chinese

In the Russian and Chinese societies people introduce themselves by saying first surname, e.g., *Wáng*, and then first name, e.g., *Lán*. This order of appearance, *Wáng Lán* 'Lán Wáng', is obligatory and it corresponds to the way in which all Chinese names are pronounced and listed at work, at school, on television, and so forth. In Great Britain and US, one introduces oneself by giving one's first name

and then surname, e.g., *Peter Jensen* or *George Robertson*. If you start with your surname, then you will repeat it when you give your first name, e.g., *The name is Bond – James Bond*. We shall argue that surname functions as (family) ground, while first name functions as figure. If this is true, Russian and Chinese have 'from-ground-to-figure' as framing strategy, whereas English and Danish have 'from-figure-to-ground' as its strategy.

How to write an address in American English and Chinese

If we turn to the ways in which envelopes are addressed in Chinese, Danish and English speaking countries, we will more or less witness exactly the same, but this time we are in a much better position, since (some of what we call) *scene* is also included. This makes address information on an envelope directly comparable to text-initial word order. Chinese envelopes always exhibit the following strict order of postal address elements: China – Province – City – Road – Person. The logic is straight-forward: the letter first reaches China, then the Province, hereafter the city, so the road and finally the person. The logic is quite another on an US envelope – it seems to be addressee oriented, which certainly tells another story, not a dynamic, but a more static one telling where one is living: Person – Road – City – Province – Country. As should be evident, everything is written in the opposite order to that of the Chinese one. While Chinese envelopes have Scene – Ground – Figure, English ones have Figure – Ground – Scene. Once again we observe that ground precedes figure in Chinese, but the opposite in English. Until quite recently Russian people used the same order as Chinese people, but a new law stopped what has been functioning for centuries. And since all envelopes are preprinted, there is actually no choice. However, if you ask them to write their address on a piece of paper, they will all do it in the old way; that is, the Chinese way of doing it.

How to present time in American English and Chinese

When we look at the way in which the same languages write the exact date and time of a past event or planned future event, the picture is the same when we consider Chinese data, but a little bit unclear, when we consider English data, because there are differences among varieties of English. In Chinese one will always start with the year, then the month and finally the date which may be followed by a

specification of the week day and a point in time: 2016 November 30 Monday noon. In other words, Chinese starts with the set '2016' and move through different subsets in order to arrive at a single element. We emphasize that this is the case for all types of oral and written discourse in Chinese – it is not only confined to official documents as it is in Sweden or to computer screens as it is in many countries. In Great Britain one will always have the order day, month, year, e.g., 30 November 2016. Here one starts with an element and ends up with the set. In US, the standard is Wednesday, November 30, 2016, but if one looks into ordinary writing you will find different ways of doing it: at noon on November 30, 2016 – at 9 a.m. on Wednesday 30 November, 2016. The important thing is, however, that we will not find the Chinese way of doing it in US or in Great Britain. And you will not find the American English or the British way of doing it in China.

5.3 *Summary*

The results stemming from eye-tracking studies of American (cf. Chua et al., 2005) and other studies (e.g., McKone et al., 2010) correspond with the word order patterns in English: the figure was fixated sooner and longer by the American graduate students when looking at a new still picture ≈ the figure appears before the ground in a text-initial utterance. And the results from Chinese students also correspond with the word order patterns in Chinese: the figure was fixated later and shorter by the Chinese graduate students when looking at a new still picture ≈ the ground appears before the figure in a text-initial utterance. American people focus on focal objects, i.e. the figure of a picture, or Caucasian Australians are characterized by having local attention, whereas Chinese people focus on contextual information, i.e. the scene and the ground, or East Asians are characterized by having global attention (cf. McKone et al., 2010). In short, it seems as if the above-mentioned results from experimental studies seem to match data from text-initial word order, the way in which people introduce themselves, the way in which they address an envelope and the way in which they treat time expressions (for more about that, see Author, 2011; 2012). Moreover, word frequency studies reveal that English mothers use more nouns which name objects, whereas Chinese mothers use more verbs that relate objects to grounds or to context. From this we cannot conclude that language determines the way we look at a picture, but we can conclude that it seems to be the case that Americans tend to look at a picture and wrap a visual scene into words in different ways compared to those of Chinese

people. It seems to be the case that English-speaking people perceive things in a local way with focus on the figure, while Chinese-speaking people perceive things in a global way with focus on the ground/scene. But we could go further. If we apply the distinction between an idea-based and image-based approach to perceiving a complex scene, it makes complete sense to argue that while an image-based approach will necessarily involve the obligatory elements of a picture, viz. a figure on a ground, an idea-based approach will necessarily mean the idea behind a picture which must be a single entity, i.e. the focal object (or figure). When dealing with categorization, we concluded that Chinese seem to have an image-based approach. When dealing with perception, we concluded that American people seem to have an idea-based form of perceiving. It may look like a paradox that categorization, a form of cognition, can be influenced by perception and perception by cognition, but it need not be so, if we acknowledge that everything has an image side as well as an idea side, which is reflected in words made up of an image-idea pair: experiences and understanding go hand in hand.

6. *Cognizing in an abstract or concrete way*

6.1 *Cognition and reasoning*

So far we have looked at potential cultural influences on primarily perceptual processes such as object categorization, perceptual judgement, and the perception of visual scenes. How, though, does culture influence more archetypal examples of higher cognition, or executive functions, such as reasoning? At least three types of reasoning have been considered in the literature on cognition and cultural differences. (Beyond reasoning, other examples of higher cognition include problem solving, judgement, and decision making.)

6.1.1 *Logical reasoning*

The first, and perhaps the archetypal, example of reasoning is logical reasoning. In studies on logical reasoning, participants are asked to assess the validity of formal arguments such as deductive arguments and syllogisms (e.g., Norenzayan, Smith, Kim, & Nisbett, 2002). The main finding is that Western participants are more likely to assess the validity of such arguments in terms of logic and internal coherence; that is, whether or not the conclusion follows logically from the premises,

irrespective of whether the conclusion or the premises are actually true. East Asian participants, on the other hand, are more likely to assess such arguments by making reference to their knowledge about external reality. For example, East Asians are more likely to accept a valid deductive argument about birds, if that argument involves the typical example of an eagle rather than the atypical example of a penguin (typicality effect; Norenzayan et al., Study 3). Similarly, East Asians are more likely to accept a valid syllogism if the conclusion is believable (belief bias; Norenzayan et al., Study 4).

6.1.2 *Causal reasoning*

A second example is how we reason about the possible causes giving rise to an observable event. One of the central findings in social psychology is the *fundamental attribution error* (FAE; Ross, 1977), where observers are more likely to attribute the causes of another person's behaviour to internal (dispositional) factors rather than external (situational) ones, even if the latter factors are greater determinants than the former. Although it was originally assumed that the FAE occurred universally, cross-cultural research suggests that Westerners are more likely to attribute the causes of events to dispositional factors than to situational ones, while East Asians are more likely to attribute the causes of another person's behaviour to situational factors than to dispositional ones (e.g., Choi, Nisbett, & Norenzayan, 1999; Norenzayan & Nisbett, 2000). This finding extends to the perception of animal behaviour (e.g., an individual fish either leading or being chased by a group of fish; Morris & Peng, 1994) and the ambiguous movements of physical objects (e.g., a stone either floating or sinking in water; Peng & Nisbett, 1997). Causal reasoning also relates to the perception of control: Westerners tend to see control as being located internally in the individual, whereas East Asians tend to see control as being located externally in the environment. (Cf. the cultural value dimension of internal/external control proposed by Trompenaars & Woolliams, 2003, pp. 91-97; 'we control nature' vs. 'nature controls us'.)

6.1.3 *Dialectical reasoning*

The third example, dialectical reasoning, is concerned with how we deal cognitively with a contradiction between two pieces of information, or a conflict between two groups of people (e.g.,

Peng & Nisbett, 1999). Dialectical reasoning relates to the Chinese notions of 'yin' and 'yang' (where opposites are seen as complementing and presupposing each other), and to the Marxian and Hegelian notions of 'thesis', 'antithesis', and 'synthesis'. The research on cultural differences suggests that Westerners are more likely to deal with contradictions by accepting one side and rejecting the other, while East Asians are more likely to see the truth or value in both points of view and to try to find some kind of 'middle way'. This basic pattern of results extends to a variety of domains including dialectical and non-dialectical proverbs, vignettes describing social conflicts, formal arguments, and scientific statements.

6.2 *Language parallels*

6.2.1 *Logical reasoning: Image-based and idea-based thinking*

In the present study English-speaking participants were contrasted with Korean ones whereby we have no indication of what Chinese participants would have done in the task involving logical reasoning. If, however, we base ourselves on identical behaviour in all other studies involving both Chinese and Korean participants and thus stipulate that Chinese people would behave in a way which is more similar to the Korean participants than to the American one, it is possible to explain this by their image-based thinking. Exactly this type of thinking will yield a typicality effect.

American students' interest in logical reasoning is also explained by their idea-based thinking, but we can be more specific than that. It appears that the various tense forms of American English involve logical reasoning. The present perfect in English involves abductive reasoning. *John has been to Paris* means that John has the psychological quality of knowing Paris (Result) because he physically spent some time there (Case) based on the fact that if you live at a certain place for some time, then it gives you a certain knowledge of that place (Rule). The same is true of *I have painted the house*: The house is painted (Result), because I painted the house (Case). Both cases involve abductive reasoning: you go back in time (Case) in order to arrive at an explanation of the present (Result) by applying a general rule. The simple past involves inductive reasoning. The utterance *He killed the cat* presents the cause "He did something" and the effect "The cat is dead" as one piece of information just as the rule "If the bean is from this bag (cause), then it is white (effect)" is condensed

into one single utterance “All beans from this bag are white”, if it is applied as many times as it is possible. The progressive aspect involves deductive reasoning. *John is scoring another goal* says that “John is producing an activity (cause)” and by applying the rule “if cause, then effect” we arrive at the conclusion “the ball is behind the line (effect)”. This is deduction. In other words, from an English point of view it is not surprising that American participants use logical reasoning – this type of thinking is imbedded in their English mother tongue (cf. Author, 2011). Note that Chinese has no tense forms at all (with the consequences that has for time reception and story-telling, cf. Lin, 2006 and Mo, 1986), and although Chinese has a form that resembles the progressive aspect in use, it doesn’t contain any reference to logic whatsoever, only to an ongoing activity based on a state consisting of a figure-ground relationship (it is composed of *zài* (cf. above) plus any activity verb).

Note also that in English one says “North, South, East, and West”, and one could not use another order. For instance, the order “East, South, West, and North” would sound illogical – nevertheless, this is the order used by Chinese: this order derives naturally from visually following the sun from sunrise to sundown. In English, North is opposed to South and East to West. They are defined in binary terms, in opposition to one another. This is another example of idea-based thinking. This is surely not the case in Chinese. Here we observe a continuum way of thinking based on perception – an order that tells a visually-based story. This is a general phenomenon and also concerns narration, as noted by Mo (1986: 308): “The events … are in the narrative mode conceived as occurring in an immediate present which (however remote the past in which they are known as to have happened) unfolds directly and without mediation before the reader’s very eyes.”

6.2.2 *Causal and dialectical reasoning*

Research on causal reasoning suggests that Westerners are more likely to locate the causes for certain actions and events *inside* the person (in the form of dispositions), whereas East Asians are more likely to locate those causes *outside* the person (in the surrounding situation). Are there any parallels for this cognitive bias in language?

The fact that English has a distinction between definite and indefinite articles, i.e. *the man* vs. *a man*), a distinction between present perfect and simple past (i.e. *has said* vs. *said*), a distinction

between *there*- and *it*-sentences and no distinction between sentences having a subject and sentences without a subject all point in the same direction: English grammar forces the speaker to communicate various pieces of information to the hearer and is not interested in the situation itself. If it were, we would not have had *there*- and *it*-sentences in which there is nothing to refer to, but, nevertheless, there must be a so-called preliminary subject in the shape of *there* or *is*. In Chinese there are many so-called subjectless sentences following the rule: if there is no figure in the picture being present on my perceptual screen, there is no subject (see Han, 2008):

(12) *hǎo* *le.*
good experience
'It is good.'

If there are some figures, you make them subject (see 13 and 14) which make them look very different from their English counterparts (cf. the translations):

(13) *yǒu* *xiē* *míngzi* *nǐ* *juéde* *hǎo.*
ground-figure some names you find good
'There are some names that you find good.'

(14) *xià* *qǐ* *yǔ* *le.*
down fall rain perf
'It has started to rain.'

An English speaker presents old pieces of information to the hearer by using the definite article, the simple past and *it*-sentences, whereas he or she presents new pieces of information to the hearer by using the indefinite article, the present perfect and *there*-sentences.

But how do we define a piece of information? A piece of information is the result arrived at by the speaker's comparing his own internal universe with that of the hearer; if there is identity, then it is old information; if there is no identity, then it is new information. The speaker's constant comparing his own mental universe with that of the hearer makes the speaker speak about reality through the hearer's mental universe and in that way it is possible to explain why American people pay attention to dispositional factors: as speakers of English they are forced to focus on the mental

universe of the hearer. If they did not, they could not know what is new or old information and this vital for the way in which one communicates in English.

One would probably think that Chinese grammar forces its speakers to communicate in the same way, but without *it*-and *there*-sentences, without definite and indefinite articles, without a distinction between present, past and future, without relative clauses and *that*-clauses it goes without saying that Chinese communicate things very differently from English speakers. Of course, as all speakers in the world, the speakers of Chinese also distinguish new and old pieces of information, but this is done by word order and as such it is governed by pragmatic rules (as it is in Russian). The distinction between new and old information is not part of Chinese grammar as it is in English, where this distinction has penetrated nominal, verbal and syntactic categories.

One may ask the following question: Is it possible to communicate other entities than pieces of information by having a specific article system, a specific tense system and a distinction between *there*-sentences and *it*-sentences? So it seems if we take a close look at the grammar of Chinese. As already mentioned above, Chinese does not know of articles, tense forms or *there*- and *it*-sentences, for that matter. Instead, they have other categories such as classifiers, a certain type of aspect and sentence final particles, such as *ba, ma, le, ne, a* and *ya* (three of them have something to do with the speaker's experience) which are not found in English at all. These grammatical categories indicate that Chinese grammar forces its speakers to communicate in a different way, although Chinese communication involves the same obligatory elements, viz. a speaker, a hearer and something to talk about, reality. Whereas in English we find a focus on the hearer, we find a focus on the situations in reality in Chinese mediated by their perception involving all senses. This focus is supported by many different pieces of evidence. First of all, Chinese is a language with serial-verb constructions. This means that Chinese reflects the order of activities appearing in reality in an iconic way. In English one says *I went to France on vacation*. You name one action, nothing more, although you actually performed many different activities between your starting point, say NY, and your end point, say Paris. In Chinese it is completely impossible to communicate the same in this abbreviated version. One may put it in several ways depending on the specific activities performed from NY to Paris, but the shortest possible version is the following, cf. (15):

(15)	*wǒ*	*zuò*	*fēijī*	*qù fǎguó*	*lǚyóu*
	I	sit down	aeroplane	go France	travelling around.

'I went (=flew) to France on vacation.'

What we witness here is an interest in the visual experience of situations having taken place and which will be taking place in that specific order they were actually produced. In other words, there is an iconic relationship between the situations referred to, the way in which they are experienced by the speaker and the order of actual verbs being used in Chinese. This does also mean that Chinese only knows of a prospective viewpoint, i.e. one always goes forward in Chinese. In English you find the prospective viewpoint (in all progressive forms) as well as the retrospective viewpoint (in all compound past forms).

The presence of a classifier in Chinese also marks a reference to external reality where we find situations, while lack of a classifier point to internal reality, i.e. the imagined world. Add to this, that subjectless sentences are used whenever there is no figure on the speaker's perceptual screen. This means that there is an object in the situation referred to that can appear as figure in the picture being presented to the hearer. In short, it seems to the case that Mandarin Chinese is a language that differs fundamentally from American English. The notion of experience, especially visual experience, seems to play a crucial role that is directly comparable to the notion of information that plays a correspondingly dominant role in English.

7. *Concluding remarks*

An increasing body of empirical evidence suggests that there are significant differences between Westerners and East Asians for a variety of psychological processes. For each of these psychological processes, we have attempted to give a comprehensive account of the parallels that exist in the native languages of the respective speakers, with focus on English as an example of a Western language, and Mandarin Chinese as an example of an East Asian language. We have found that the language parallels in question are numerous, systematic, and seem to occur without exception. This means that we find two important areas of evidence which point in the same direction. External pieces of empirical evidence from a wide range of cognitive areas and internal pieces of evidence from a wide

range of linguistic areas demonstrate important cultural cognitive differences between American English speaking persons and Chinese speaking persons. The American English linguacultural universe seems to be very different from the Mandarin Chinese linguacultural universe. The effect of including linguistic evidence should be obvious. Whereas the examined studies involve an extremely small number of members of the respective speech communities, the linguistic analyses of English and Chinese in principle involve all members of the two speech communities. This is the case because language is a conventional semiotic tool that is negotiated on a daily basis and is used to communicate among all members of a speech community without exception. Because the linguistic differences reflect the cultural cognitive differences, the empirical evidence is so strong that it is not possible anymore to ignore them by arguing that the evidence is too weak or vague. But we emphasize that this is in no way tantamount to saying that language determines culture and cognition. We would argue that the problem is far more complex than that and therefore also far more complex than traditionally believed (cf. the Sapir-Whorf hypothesis, Whorf, 1956; Sapir, 1921; for an overview, see Sharifian 2017a). It stresses, however, the importance of studying and discussing the question of the correlation between the various factors involved. If we take perception, cognition, action/interaction, and communication as a *tertium comparationis*, it seems to be more or less obvious that both culture and language must be influential. The question, however, remains: Does language only reflect cultural cognition and function as a store of cultural heritage, or do some areas of language also have an active, and not only a retroactive, influence on cultural cognition?

In our search for a suitable general term for *image-based* and *idea-based* we have at our disposal a range of possibilities: image schema (Johnson 1987; Lakoff 1987), cultural schema (Nishida 1999), cultural conceptualization (Sharifian 2017a; 2017b), and mental models (Johnson-Laird 1983; 1996) to list some of the most important and obvious notions to apply. Our problem is, however, that if we argued that in the case of American and Chinese linguacultures we were dealing with different image schemas, different cultural schemas, different cultural conceptualizations or different mental models, we would treat *image-based* and *idea-based* as something which differed fundamentally from one another without implying that the two notions are complementary. They are

only seemingly mutually exclusive, because you cannot apply both approaches at the same time, but have to choose between them.

All people irrespective of culture constantly experience (by means of images) and understand (by means of ideas) things and situations when being confronted with them. And all languages make use of words and utterances when communicating about experiences and understandings. In all human beings an experience goes hand in hand with understanding and in all languages it is impossible to separate the image side, i.e. the experiential part of an expression unit linked to all the human senses (the body), from its idea side, i.e. the knowledge part of an expression unit linked to the human mind/ brain. As shown in Bentsen (to appear), in any word for "bread" in any language it will not be possible to separate the images deriving from seeing, smelling, tasting and feeling a piece of bread from the ideas that tell you that it is a piece of artefact made by people to serve as a kind of food that is made out of some sort of flour and some sort of liquid, water or milk or even eggs, which is then baked in an oven. In all linguacultures the word for "bread" will incorporate certain images and certain ideas, but not exactly the same, because a prototypical bread of a French person would differ from the prototypical bread of an Italian that would differ from a prototypical bread of a German person.

Our problem is to find a term that itself presupposes that people share the same way of processing external stimuli and that languages share the same way of representing the final outcome of this mental processing, but, nevertheless, differ in their focus. We will suggest the notion of *perspective* (cf. Verhagen 2007), because this term presupposes that something is shared while at the same time allowing for focusing on different parts of what is common to all. In short, it is our hypothesis that American and Chinese participants have different approaches to solving all kinds of problems in empirical tests, because they have different perspectives: American participant use an idea-based perspective, while the Chinese use an image-based perspective. American and Chinese participants may have the same experiences and the same understanding of a problem in a task, but due to different perspectives they will, nevertheless, arrive at different results as research shows. Likewise, it is our hypothesis that American English and Mandarin Chinese differ from one another because different perspectives are involved: American English focuses on new and old information in

its grammatical structure (by having a nominal distinction between the definite and indefinite article, by having a verbal distinction between present perfect and simple past, and by having a syntactic distinction between *there*- and *it*-sentences) because it is idea-based in its perspective, whereas Mandarin Chinese (that does not know of these grammatical distinctions) focuses on the speaker's experiences in its grammatical structure (by having different classifiers before different nouns, by having a series of verbs instead of one verb in each sentence, and by having so-called sentence final particles) because it is image-based in its perspective.

The image-based perspective applied by Chinese participants in the empirical tests examined by us and the image-based perspective found in their mother tongue, i.e. Mandarin Chinese, and the idea-based perspective applied by the American participants in the same tests and the idea-based perspective found in American English seem to be the effect of important differences in cultural preferences for focusing on one thing at the dispense of another thing although both are equally present. It seems to be the case that members of the Mandarin Chinese speech community have a specific interest in speaking about their experiences of objects and situations in external reality, whereas members of the American English speech community have a specific interest in speaking about their understanding of these experiences. Our memory is composed of sensory memory (images), working memory (ideas) and long-term memory where the outcome of the input and the intake is found (cf. Atkinson & Schiffrin 1968; see also Cowan 2008; Schweppe & Rummer 2014). In short, our long-term memory consists of (pieces of) still pictures and motion pictures accompanied by our understanding of those objects and situations that gave rise to these experiences. The preference for focusing on images or ideas cannot be explained by language itself, but it is maintained by language use. We find it likely that the two very different writing systems are the results of applying two different perspectives to the content of an expression unit: the Western writing system based on an alphabetic principle reflects the idea-based perspective, while the East Asian writing system based on a logographic principle reflects the image-based perspective. If this is true, then there is a crucial retroactive influence by language on culture.

References

Aikhenvald, A. Y. (2000). *Classifiers: A typology of noun categorization devices*. Oxford: Oxford University Press.

Allan, K. (1977). Classifiers. *Language*, *53*, 285-311.

Atkinson, R.C. & Schiffrin, J.C. (1968). Human memory: a proposed system and its control processes. In K.W. Spence & j.t. Spence (Eds.), *The psychology of learning and motivation*, 89-195. New York: Academic Press.

Baron, I. & Herslund, M. (2005). Langues endocentriques et langues exocentriques. Approche typologique du danois, du français et de l'anglais. In M. Herslund & I. Baron (eds.), *Le génie de la langue française. Perspectives typologiques et contrastives*. [Special edition]. *Langue française 145*, 35-53.

Barsalou, L.W. (1999). Perceptual symbol systems. *Behavioral Brain Sciences*, 22, 577-660.

Bentsen, S.E. (to appear). *The comprehension of English texts by non-native speakers of English from China, Japan and Russia.*

Boduroglu, A., Shah, P., & Nisbett, R. E. (2009). Cultural differences in allocation of attention in visual information processing. *Journal of Cross-Cultural Psychology*, *40*, 349-360.

Boland, J. E., Chua, H. F., & Nisbett, R. E. (2008). How we see it: Culturally different eye movement patterns over visual scenes. In K. Rayner, D. Shen, X. Bai, & G. Yan (eds.), *Cognitive and cultural influences on eye movements* (pp. 363-378). Tianjin, CH: Tianjin People's Publishing House.

Castelhano, M. S. & Rayner, K. (2008). Eye movements during reading, visual search, and scene perception: An overview. In K. Rayner, D. Shen, X. Bai, & G. Yan (eds.), *Cognitive and cultural influences on eye movements* (pp. 3-33). Tianjin, CH: Tianjin People's Publishing House.

Chhokar, J. S., Brodbeck, F. C., & House, R. J. (eds.) (2008). *Culture and leadership across the world. The GLOBE book of in-depth studies of 25 societies*. New York/Abingdon: Lawrence Erlbaum Associates.

Chiu, L. H. (1972). A cross-cultural comparison of cognitive styles in Chinese and American children. *International Journal of Psychology*, *7*, 235-242.

Choi, I., Nisbett, R. E., & Norenzayan, A. (1999). Causal attribution across cultures: Variation and universality. *Psychological Bulletin*, *125*, 47-63.

Chomsky, N. (1975). *Reflections on language*. New York: Pantheon.

Chua, H. F., Boland, J. E., & Nisbett, R. E. (2005). Cultural variation in eye movements during scene perception. *Proceedings of the National Academy of Sciences USA*, *102*, 12629-12633.

Cowan, N. (2008). What are the differences between long-term, short-term, and working memory? *Progress in Brain Research* 169, 323-338.

Durst-Andersen, P. (2009). The grammar of linguistic semiotics. Reading Peirce in a modern linguistic light. *Cybernetics & Human Knowing*, 16, 38-79.

Durst-Andersen, P. (2011). *Linguistic supertypes. A cognitive-semiotic theory of human communication.* Berlin/New York: De Gruyter Mouton.

Durst-Andersen, P. (2012). What languages tell us about the structure of the human mind. *Cognitive Computation*, *4*, 82-97.

Evans, K., Rotello, C. M., Li, X., & Rayner, K. (2009). Scene perception and memory revealed by eye movements and receiver-operating characteristic analyses: Does a cultural difference truly exist? *The Quarterly Journal of Experimental Psychology*, *62*, 276-285.

Fernald, A. & Morikawa, H. (1993). Common themes and cultural variations in Japanese and American mothers' speech to infants. *Child Development*, *64*, 637-656.

Gao, D. G. & Kao, H. S. R. (2002). Psycho-geometric analysis of commonly used Chinese characters. In H. S. R. Kao & D. G. Gao (eds.), *Cognitive neuroscience studies of the Chinese language* (pp. 195-206). Hong Kong: Hong Kong University Press.

Gerlach, C. (2008). *Visual object recognition and category specificity.* Post-doctoral dissertation. Copenhagen: Copenhagen University Press.

Gerlach, C. (2009). Category-specificity in visual object recognition. *Cognition*, *111*, 281-301.

Hall, E. T. (1976/1997). *Beyond culture.* New York: Anchor Books.

Han, X. (2009). Maybe There Are No Subject-Predicate Sentences in Chinese. *Dao*, *8*(3), 277-287.

Harnad, S. (1990). The symbol grounding problem. *Physica D*, 42, 335-346.

Hedden, T., Ketay, S., Aron, A., Markus, H. R., & Gabrieli, J. D. E. (2008). Cultural influences on neural substrates of attentional control. *Psychological Science, 19*(1), 12-17.

Herslund, M. & Baron, I. (2003). Language as world view. Endocentric and exocentric representations of reality. In I. Baron (ed.), *Language and culture*. [=*Copenhagen Studies in Languages 29*], 29-42. Copenhagen: Samfundslitteratur.

Hofstede, G. (1980). *Culture's consequences: International differences in work-related values*. Beverly Hills, CA: Sage Publications, Inc.

Hofstede, G. (1991). *Cultures and organizations: Software of the mind*. London, UK: McGraw-Hill.

House, R. J., Hanges, P. J., Javidan, JM., Dorfman, P. W., & Gupta, V. (Eds.) (2004). *Culture, leadership, and organizations. The GLOBE study of 62 societies*. Thousand Oaks, CA: Sage Publications, Inc.

Hu, J., & Pan, H. (2008). Focus and the basic function of Chinese existential you-sentences. In Comorovski, I., & Von Heusinger, K. (Eds.) *Existence: Semantics and syntax* (Vol. 84). (pp. 133-145). Dordrecht: Springer Science & Business Media.

Ito, K., Masuda, T., & Hioki, K. (2012). Affective information in context and judgment of facial expression: Cultural similarities and variations in context effects between North Americans and East Asians. *Journal of Cross-Cultural Psychology*, *43*(3), 429-445.

Ji, L., Peng, K. & Nisbett, R. E. (2000). Culture, control, and perception of relationships in the environment. *Journal of Personality and Social Psychology*, *78*, 943-955.

Johnson, M. (1987). *The body in the mind. The bodily basis of meaning, imagination, reasoning.* Chicago: The University of Chicago Press.

Johnson-Laird, Ph. N. (1983). *Mental models: Towards a cognitive science of language, inference, and consciousness.* Cambridge: Cambridge University Press.

Johnson-Laird, Ph. N. (1996). Space to think. In P. Bloom et al. (eds.), *Language and space*, 437-462. Cambridge, MA: The MIT Press.

Kay, P., & Kempton, W. (1984). What is the Sapir-Whorf hypothesis? *American Anthropologist*, *86*, 65-79.

Kitayama, S., Duffy, S., Kawamura, T., & Larsen, J. T. (2003). Perceiving an object and its context in different cultures: A cultural look at New Look. *Psychological Science*, *14*, 201-205.

Klein, W., Li, P. & Hendriks, H. (2000). Aspect and assertion in Mandarin Chinese. *Natural Language and Linguistic Theory*, 18, 723-770.

Lakoff, G. (1987). *Women, fire, and dangerous things: What categories reveal about the mind.* Chicago: The University of Chicago Press.

Li, X., Williams, C. C., Cave, K. R., & Well, A. R., & Rayner, K. (2008). Eye movements, individual differences, and cultural effects. In K. Rayner, D. Shen, X. Bai, & G. Yan (eds.), *Cognitive and cultural influences on eye movements* (pp. 379-393). Tianjin, CH: Tianjin People's Publishing House.

Li, C. N. & Thompson, S. (1981). *Mandarin Chinese: A functional reference grammar*. Berkeley & Los Angeles: University of California Press.

Lin, J-W. (2003). Aspectual selection and negation in Mandarin Chinese. *Linguistics*, 41, 425-459.

Lin, J-W. (2006). Time in a language without tense: the case of Chinese. *Journal of semantics*, 23, 1-53.

Liu, W., Inhoff, A. W., Ye, Y., & Wu, C. (2002). Use of parafoveally visible characters during the reading of Chinese sentences. *Journal of Experimental Psychology: Human Perception and Performance*, *28*, 1213-1227.

Masuda, T., Ellsworth, P. C., Mesquita, B., Leu, J., Tanida, S., & Van de Veerdonk, E. V. (2008). Placing the face in context: Cultural differences in the perception of facial emotion. *Journal of Personality and Social Psychology*, *94*(3), 365-381.

Masuda, T., Gonzalez, R., Kwan, L., & Nisbett, R. E. (2008). Culture and aesthetic preference: Comparing the attention to context of East Asians and Americans. *Personality and Social Psychology Bulletin*, *34*, 1260-1275.

Masuda, T., Ishii, K., & Kimura, J. (2016). When does the culturally dominant mode of attention appear or disappear? Comparing patterns of eye movement during the visual flicker task between European Canadians and Japanese. *Journal of Cross-Cultural Psychology*. DOI: 10.1177/0022022116653830.

Masuda, T. & Nisbett, R. E. (2001). Attending holistically vs. analytically: Comparing the context sensitivity of Japanese and Americans. *Journal of Personality and Social Psychology*, *81*, 922-934.

Masuda, T. & Nisbett, R. E. (2006). Culture and change blindness. *Cognitive Science*, *30*, 381-399.

McKone, E., Davies, A. A., Fernando, D., Aalders, R., Leung, H., Wickramariyaratne, T., & Platow, M. J. (2010). Asia has the global advantage: Race and visual attention. *Vision Research*, *50*, 1540-1549.

Mo, T. (1986). *An insular possession*. London: Chatto & Windus.

Morris, M. W. & Peng, K. (1994). Culture and cause: American and Chinese attributions for social and physical events. *Journal of Personality and Social Psychology*, *67*, 949-971.

Nand, K., Masuda, T., Senzaki, S., & Ishii, K. (2014). Examining cultural drifts in artworks through history and development: cultural comparisons between Japanese and western landscape paintings and drawings. *Frontiers in Psychology*, *5*, Article 1041. DOI: 10.3389/fpsyg.2014.01041.

Nisbett, R. E. (2003). *The geography of thought: How Asians and Westerners think differently … and why*. London: Nicholas Brealey Publishing.

Nisbett, R. E. & Masuda, T. (2003). Culture and point of view. *Proceedings of the National Academy of Sciences USA*, *100*, 11163-11170.

Nisbett, R. E. & Masuda, T. (2006). The influence of culture: Holistic versus analytic perception. *Trends in Cognitive Sciences*, *9*, 467-473.

Nisbett, R. E., Peng, K., Choi, I., & Norenzayan, A. (2001). Culture and systems of thought: Holistic versus analytic cognition. *Psychological Review*, *108*, 291-310.

Nishida, H. (1999). Cultural Schema Theory: In W.B. Gudykunst (ed.), *Theorizing About Intercultural Communication*, 401-418. Thousand Oaks, CA: Sage Publications, Inc.

Norenzayan, A. & Nisbett, R. E. (2000). Culture and causal cognition. *Current Directions in Psychological Science*, *9*, 132-135.

Norenzayan, A., Smith, E. E., Kim, B. J., & Nisbett, R. E. (2002). Cultural preferences for formal versus intuitive reasoning. *Cognitive Science*, *26*, 653-684.

Peirce, C. S. 1998. *The essential Peirce: Selected philosophical writings Volume 2 (1893-1913)*, ed. by the Peirce Edition Project. Bloomington, IN: Indiana University Press.

Peng, K. & Nisbett, R. E. (1997). Cross-cultural similarities and differences in the understanding of physical causality. In M. Shield (ed.), *Proceedings of conference on culture and science*. Kentucky, KY: Kentucky State University Press.

Peng, K. & Nisbett, R. E. (1999). Culture, dialectics, and reasoning about contradiction. *American Psychologist*, *54*, 741-754.

Peyraube, A. (2006). Motion verbs in Chinese. A diachronic study of directional complements. In Maya Hickmann & Stéphane Robert (eds.), *Space in languages. Linguistic systems and cognitive categories* (121–135). Amsterdam & Philadelphia: John Benjamins.

Rayner, K., Li, X., Williams, C. C., Cave, K. R., & Well, A. R. (2007). Eye movements during information processing tasks: Individual differences and cultural effects. *Vision Research*, *47*, 2714-2726.

Rensink, R. A., O'Regan, J. K., & Clark, J. J. (1997). To see or not to see: The need for attention to perceive changes in scenes. *Psychological Science*, *8*, 368-373.

Rizzolatti, G., Riggio, L., Dascola, I., & Umiltá, C. (1987). Reorienting attention across the horizontal and vertical meridians: Evidence in favor of a premotor theory of attention. *Neuropsychologia*, *25*, 31-40.

Rubin, E. (1915). Synsoplevede Figurer: Studier i psykologisk Analyse. Første Del [Visually experienced figures: Studies in psychological analysis. Part one]. Copenhagen and Christiania: Gyldendalske Boghandel, Nordisk Forlag.

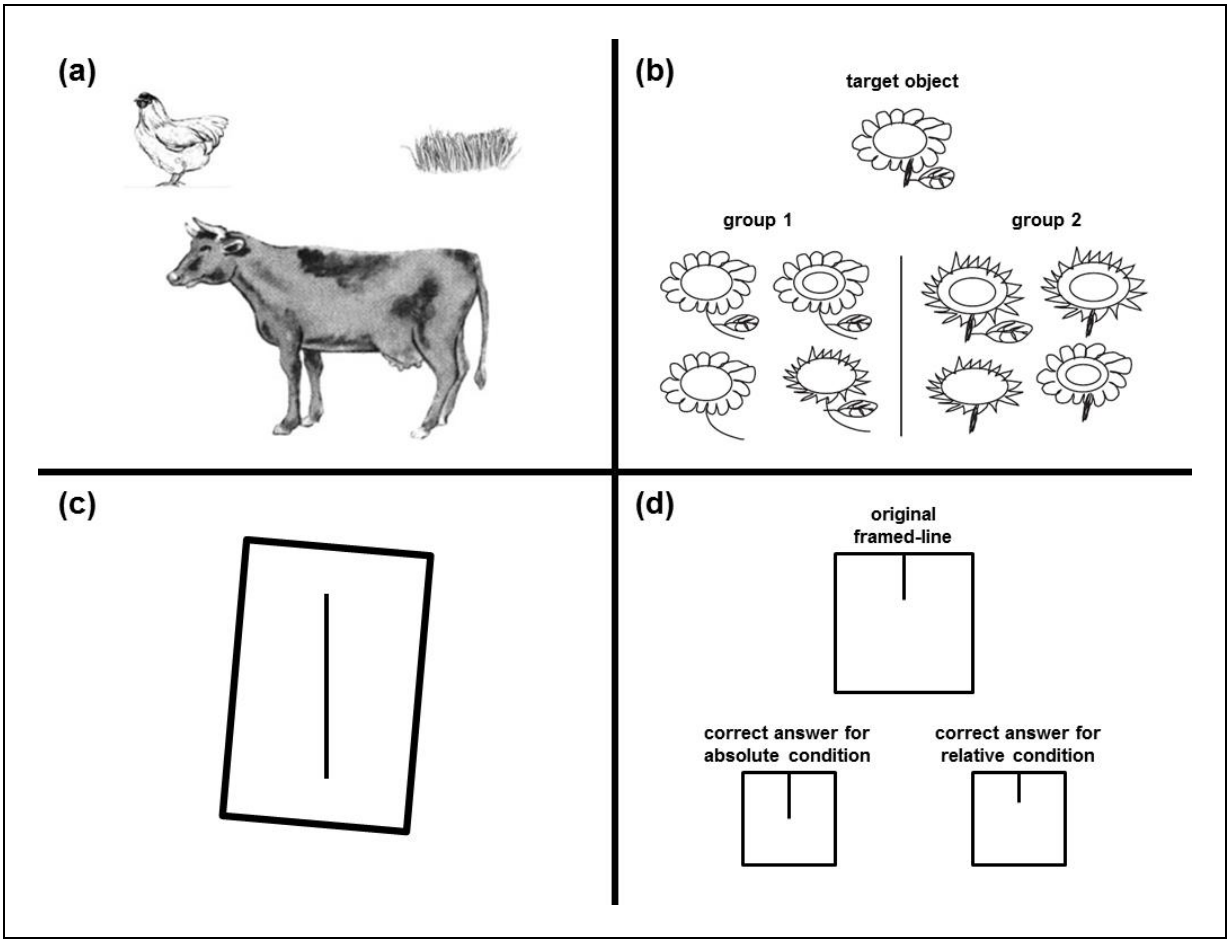
Sapir, E. (1921). *Language. An introduction to the study of speech*. New York, NY: Harcourt, Brace & World, Inc.

Saussure, F. de (1916). *Cours de linguistique générale.* Paris: Payot.

Schweppe, J. & Rummer, R. (2014). Attention, working memory, and long-term memory in multimedia learning: An integrated perspective based on process models of working memory. *Education and Psychology Review* 26 (2), 285-306.

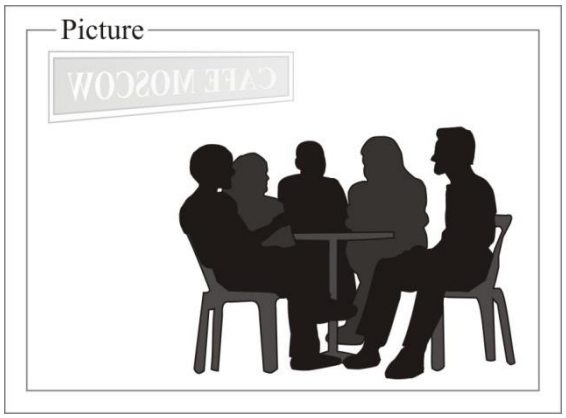
Senzaki, S., Masuda, T., & Ishii, K. (2014). When is perception top-down and when is it not? Culture, narrative, and attention. *Cognitive Science*, *38*, 1493-1506.
Senzaki, S., Masuda, T., & Nand, K. (2014). Holistic versus analytic expressions in artworks: Cross-cultural differences and similarities in drawings and collages by Canadian and Japanese school-age children. *Journal of Cross-Cultural Psychology*, *45*(8), 1297-1316.
Sharifian, F. (2017a). Cultural linguistics and linguistic relativity. *Language Sciences*, 59, 83-92.
Sharifian, F. (ed.) (2017b). *Cultural conceptualizations and language.* Amsterdam/Philadelphia: John Benjamins.
Simons, D. J. (2000). Current approaches to change blindness. *Visual Cognition*, *7*(1/2/3), 1-15.
Tardif, T., Gelman, S. A., & Xu, F. (1999). Putting the "noun bias" in context: A comparison of English and Mandarin. *Child Development*, *70*, 620-635.
Ting-Toomey, S. (1999). *Communicating across cultures.* London: The Guilford Press.
Trompenaars, F., & Woolliams, P. (2003). *Business across cultures*. West Sussex, UK: Capstone Publishing.
Tsang, Y. K. & Chen, H. C. (2008). Eye movements in reading Chinese. In K. Rayner, D. Shen, X. Bai, & G. Yan (eds.), *Cognitive and cultural influences on eye movements* (pp. 237-254). Tianjin, CH: Tianjin People's Publishing House.
Unsworth, S. J., Sears, C. R., & Pexman, P. M. (2005). Cultural influences on categorization processes. *Journal of Cross-Cultural Psychology*, *36*(6), 662-688.
Verhagen, A. (2007). Construal and perspectivisation. In D. Geeraerts & H. Cuyckens (eds.), *Handbook of Cognitive Linguistics*. Oxford: Oxford University Press.
Whorf, B. L. (1956). *Language, thought, and reality: Selected writings*. Cambridge, MA: MIT Press.
Witkin, H. A. (1967). A cognitive-style approach to cross-cultural research. *International Journal of Psychology*, *2*, 233-250.
Wu, J-S. (2005). The semantics of the perfective *le* and its context-dependency: an SDRT approach. *Journal of East Asian linguistics*, 14, 299-336.
Xiao, Z. & McEnery, A. (2004). *Aspect in Mandarin Chinese: a corpus-based study*. Amsterdam & Philadelphia: John Benjamins.
Zhong, C. H. E. N. (2007). The distributional sequence and motivation of *lai* and *qu* as complex directional complement [J]. *Contemporary Linguistics*, *1*, 004.
Yip, P-C. (2000). *The Chinese lexicon. A comprehensive survey.* London & New York: Routledge.

Figure 1



Caption: Examples of the experimental paradigms used for testing cultural differences in object categorizations and perceptual judgements: (a) object categorization task with three target objects (from Chiu, 1972); (b) object categorization task with one target object and two groups (from Norenzayan et al., 2002); (c) the Rod-and-Frame Test (RFT); and (d) the Framed-Line Test (FLT).

Figure 2



Caption: A picture of five people sitting around a table in a small café.