

# Why are you amused: Unveiling multimodal humour from the prototype theoretical perspective

**Qiaoyun Chen**

College of Foreign Languages and Cultures, Xiamen University  
[chenchen.cqyfj@163.com](mailto:chenchen.cqyfj@163.com)

**Guiying Jiang**

College of Foreign Languages and Cultures, Xiamen University  
[gyjiang118@126.com](mailto:gyjiang118@126.com)

## Abstract

*This paper looks at multimodal humour through the lens of prototype theory in the framework of conventional incongruity theory of humour, aiming for a unified linguistic and semiotic approach to humour. From this perspective, humour can be achieved through the following three aspects of linguistic and non-linguistic categories: 1) prototypicality versus non-prototypicality of category members; 2) the family resemblance shared by category members; 3) vague inter-categorical boundary. The cognitive mechanisms behind this type of multimodal humour and its comprehension are discussed. The intermodal relationships involved are examined and categorised into two major types: complementary and non-complementary ones.*

*Keywords: multimodal humour; incongruity theory of humour; prototype theory; Cognitive Linguistics.*

## 1. Introduction

In recent decades, the surging multimodal humour has paralleled the volume of traditional verbal humour, but the scale of relevant studies does not seem to match its development. What is meant by multimodal humour is the type of humour which recruits more than one semiotic code functioning synchronously in creating amusement. Typical cases include memes, cartoons, comics, sitcoms and the like that involve verbal-visual, audio-visual or all three elements. Thanks to modern multimedia technology, multimodal approach has proliferated and pervaded humorous communications nowadays, fostering increasing theoretical and empirical interests in this regard. Mechanisms for humour, particularly linguistic humour, have been extensively explored and various humour theories have been formulated from multiple perspectives. We contextualise the present study in Cognitive Linguistics (CL

henceforth) and the incongruity theory of humour for two reasons. First, in light of the cognitive orientation manifested by considerable humour research, CL and humour studies are presumably highly compatible and complementary with each other, which promotes the exchange of insights between the two (Brône et al. 2006: 205). Recurrent in most cognitive approaches to humour, the element of incongruity (Hull et al. 2017: 486) is postulated in our argument to be a prerequisite for humour. Second, theoretical constructs in CL, such as conceptual blending, metaphor, metonymy, and frame-shifting have shed new light on humour research. However, prototype theory, a theory concerning the degree of salience and typicality which we regard as an inherent source of incongruity and therefore an important inducer of humour, has not been sufficiently investigated as to how it can possibly help create the humorous effect. This article seeks to examine how theories of verbal humour fit into the category of multimodal humour and account for multimodal humour from the prototype theoretical perspective, with the incongruity theory serving as the lynchpin.

The article is structured as follows: in section 2, a cursory review is given with regard to general CL endeavours in humour and prototype theory, followed by a discussion about the conjunction of the incongruity theory of humour and prototype theory. Section 3 surveys how the major principles of prototype theory can contribute to multimodal humour in the cases of linguistic and non-linguistic categories. Intermodal relationships in humour are interpreted in section 4, while section 5 presents the conclusions.

## **2. Literature review**

### **2.1. CL endeavours in humour research**

The studies of linguistic humour since time immemorial began to take a cognitive turn in the 1980s with Raskin's seminal work on the semantic mechanisms of humour (Raskin 1985). The years since then have witnessed a steady increase in humour-related publications with cognitive orientation, a trend which has sustained to this day. According to Attardo (2002: 231), linguists who study humour are actually engaged in doing cognitive stylistics all along. Close ties between CL and humour study are delineated in Brône et al. (2006), which points to the great potential of interdisciplinary research. In what follows, we will review how some CL constructs are applied to humour research by reporting on some notable endeavours. Firstly, as is argued in Brône et al (2006: 206), if construal operations (e.g., attention and salience, comparison, perspective and viewpoint) are indeed basic-level semantic-conceptual mechanisms, they can account not only for conventional language use, but also the semantic construal of marked cases of language use like humour. In the same vein, Brône & Feyaerts (2004), by assessing the Semantic Script Theory of Humor (SSTH henceforth) and the General Theory of Verbal Humour (GTVH henceforth) within CL framework, put forward a prototype model of construal operations for humour, in which incongruity and its resolution are taken as two perspectives of the same cognitive construal. Secondly, attempts have been made to apply the prototype theory in humour research. Among others, Ferro-Luzzi (1990) proposes a polythetic-prototype approach to the conception of humour and typology of humour theories, aiming to address the inadequacy of humour theories by exploiting the principles of family resemblance and vague boundary. The role of cognitive factors in children's humour comprehension is investigated by McGhee & Panoutsopoulou (1990), who conclude that the intersection of categories and divergent thinking skills are the key contributing cognitive factors. Thirdly, as for the interaction between metaphor and humour, Kang (2016) highlights the lexical level of verbal humour, an aspect that is largely overlooked, by focusing on humorous metaphor-based nominal compounds that are marked as "humorous" in German

dictionaries. It is concluded that humour in metaphor-based nominal compounds is produced by recognition and appropriate resolution of the incongruity between the source and the target domains of metaphor on the one hand, and the literal and the figurative meanings of the compound on the other (Kang 2016: 376). Apart from verbal metaphor, visual metaphor is attested to be one of the most common humour mechanisms in cartoon drawing (Templin 1999: 21; Edwards 2001: 2142; El Refaie 2003; Plumb 2004; Tsakona 2009). The elements come from different domains and there is a likelihood that they may be incongruous to one another and thus result in the creation of humorous effect (Tsakona 2009: 1180). Moreover, more CL underpinnings such as conceptual metonymy theory (e.g., Tabacaru & Feyaerts 2016), salience and conceptual blending theory have also found application in revealing humour mechanisms (see Brône et al. 2006 for a more detailed discussion). In a nutshell, while CL opens up new perspectives for humour research, the inclusion of humour as an object of CL studies has also benefited the discipline.

As far as multimodal humour is concerned, it has been dealt with from various perspectives; to name just a few, the cognitive (Brône & Feyaerts 2003; 2004; Brône et al. 2006, etc.) and semiotic (Carroll 1996; Nikolajeva & Scott 2000; El Refaie 2003; 2004, etc.) points of view, or a combination of the two (Tsakona 2009). However, prototype theory has not been utilised to explain the multimodal humour.

To sum up, the role played by the principles of prototype theory in humour is apparently underexplored vis-à-vis other CL constructs. And in view of the status quo of multimodal humour research, we intend to approach this humorous genre from the prototype theoretical perspective by looking at how principles of the theory are capable of generating humour in language and non-language realms, drawing on various CL concepts such as construction. Our focus is on humour encoded in text and image modalities, a common form of multimodal humour.

## **2.2. Prototype theory**

This section provides a brief introduction of prototype theory. Studies on human categorisation underwent a radical change with the work of Eleanor Rosch and her colleagues in the 1970s. Rosch (1978: 27-28) proposed two general and basic principles for the formation of categories, i.e. cognitive economy and perceived world structure. The former principle influences the abstraction level of categories formed in a culture and the latter the internal structure of the categories once they are formed. Accordingly, category systems are conceived of as having two dimensions, that is, the vertical one that concerns the level of inclusiveness of a category and the horizontal one which relates to the segmentation of categories at the same level of inclusiveness. We propose a simplified illustration of Rosch's categorisation system as follows:

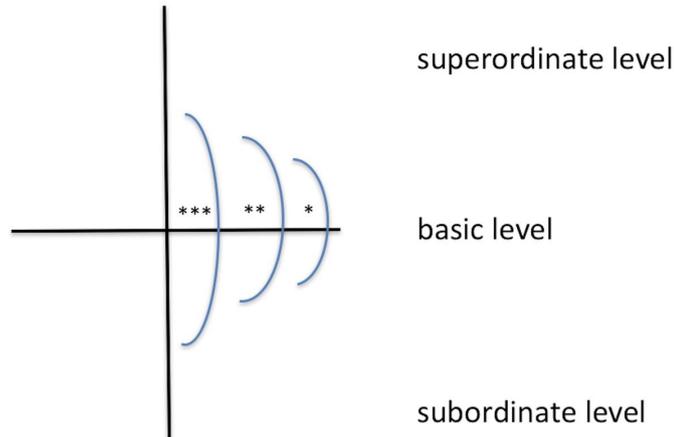


Figure 1. Illustration of Rosch's categorisation system (the asterisk roughly signals the number of attributes of category members)

Vertically, in terms of level of inclusiveness, Rosch & Mervis (1975), and Rosch et al. (1976) stratify categories into the superordinate level, the basic level and the subordinate level, forming a cline such as “living thing-animal-mammal-dog-poodle”. In this dimension, the higher up the axis, the more inclusive a category will be and vice versa. In between the poles there is a basic level which is “the most inclusive (abstract) level at which the categories can mirror the structure of attributes perceived in the world” (Rosch 1978: 30). Prototype theory resides in the horizontal axis and concerns the internal structure of categories. It was formulated by Rosch and her colleagues in the 1970s as an alternative to the traditional classical theory of human categorisation and got enriched by Lakoff (1987), Geeraerts (1989; 1997), Dirven & Verspoor (1998) and Taylor (2003). Core elements of prototypical categorisation are well captured by Ungerer & Schmid (1996: 39) as the notion of prototype, category membership and typicality, attributes, family resemblance and gestalt. Essentially, the prototype of a category is the members that “contain the attributes most representative of items inside and least representative of items outside the category” (Rosch 1978: 30). As gestalt concepts, prototypical categories tend to be defined by virtue of prototypes or prototypical instances, often exhibit typicality effect and have fuzzy boundaries. Category members, rather than sharing the same discrete attributes, are linked by family resemblance and exhibit a radial structure as illustrated in the horizontal axis of Figure 1 (the asterisk roughly signals the number of attributes of members and there is a diminishing trend toward the periphery of the categorial structure). In section 3, we will proceed to argue that the two dimensions of categorisation can contribute to humorous effect, especially the prototype theoretical dimension.

### 2.3. When the incongruity theory of humour meets prototype theory

According to Attardo (1994: 47), modern humour theories are categorised into three groups: the cognitive (incongruity), social (hostility), linguistic (SSTH and GTVH) and psychoanalytical (release) approaches. Most of these theories are propounded for verbal humour and later adapted to accommodate multimodal humour. Among these propositions, incongruity has long and widely been recognised as a necessary condition for humour (Deckers 1993). As Ritchie (2004: 46) puts it, “If there is one generalisation that can be extracted from the literature about humour, it is that humour involves incongruity”. Specifically, the incongruity theory holds that the perception of incongruity leads to humour

when two normally disparate concepts, schemata, or meanings are juxtaposed in a surprising or unexpected manner (Attardo & Raskin 1991; Coulson 2001; Ruch 2008). It has received considerable empirical support (Deckers & Buttram 1990; Hull et al. 2017; etc.). Suls (1972: 85-89) streamlines the humour appreciation process in a three-stage information processing model: a setup stage where a prediction is made for comparison with the most recent text input, an incongruity stage where the input does not match the prediction, and a resolution stage where one engages in figuring out how the punchline is congruent with the preceding text. Nevertheless, it is argued that resolution of the incongruity is not entailed in all types of humour. Therefore, a distinction should be made between humour types consisting of incongruity alone (e.g. nonsense humour) and humour with incongruity and resolution (Forabosco 1992; Attardo 1994). A critical concept in the incongruity theory of humour is the schema which is “a dynamic mental representation that enables us to build mental models of the world” (Martin 2007: 85) and “is formed on the basis of past experience with objects, scenes, or events and consists of a set of (usually unconscious) expectations about what things look like and/or the order in which they occur” (Mandler 1979: 263). In CL, it is either used interchangeably with such concepts as frame and script or distinguished from them. They will be used interchangeably below. This notion has played a crucial role in illuminating the nature of incongruity in humour and underlines SSTH and GTVH. Several schema-based theories of verbal humour have been proposed to decode the process of humour perception (Norrick 1986; Raskin 1985; Wyer & Collins 1992, etc.). Basically they all suggest that a schema is activated at the setup stage in order to help us understand the incoming information; at the incongruity stage, this schema is confronted with a lexical cue in the text, i.e., the punchline that appears incompatible with the activated schema and then a switch of schema is initiated—a second, backgrounded schema is stimulated to make better sense. Martin (2007: 87) states that the two schemas are activated simultaneously for the reason that the second schema does not replace the first one completely; that the essence of humorous incongruity rests upon this simultaneous activation of two incompatible schemas, which is experienced as enjoyable and amusing. We assume that in the case of multimodal humour, the triggers to launch the operations in the setup and incongruity stages are not confined to the single verbal code, but are instead distributed among multiple modalities.

Categorisation is constantly taking place in our mind, either consciously or unconsciously, so that we can make sense of the world. Considering the hierarchical organisation of categories (Figure 1), it is argued that vertically, the basic level, which is cognitively and linguistically more salient, is the direct and fundamental starting point where we depart in our journey of understanding and learning about the entities in the world (Wang 2006: 139). Horizontally, while prototypes or prototypical instances are highly evocative of their category, the reverse also holds true. In short, it is sensible to say that the default value of categorisation in our communication is the prototypes or more prototypical instances of basic-level categories. In other words, they are what we usually expect of categorisation in our communication. When this expectation is not met, for one reason or another, an incongruity might arise and lead to humour, as will be elaborated in the following section.

### **3. Interpretation of multimodal humour from the prototype theoretical perspective**

This section aims to explore how humour is related with linguistic and non-linguistic prototypical categories and the issue concerning the comprehension of prototypicality-based humour. It is necessary to clarify the scope of prototype in our study in the first place. Rosch and her colleagues' experiments (Rosch et al. 1976) on categorisation concentrate on





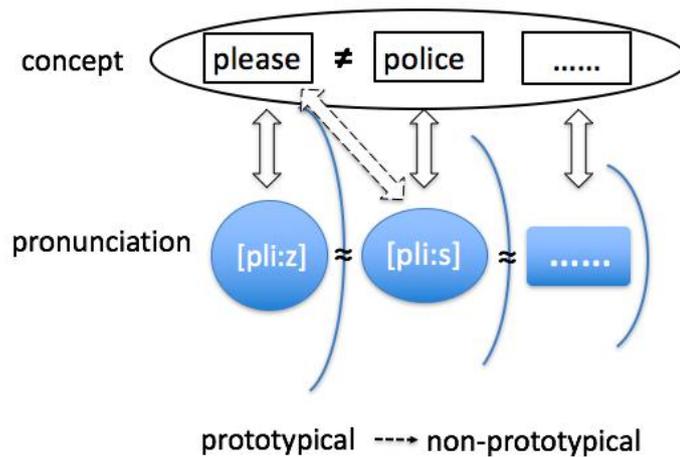


Figure 4. A specific illustration of the incongruity caused by phonological non-prototypicality  
( $\neq$ : incongruity;  $\approx$ : family resemblance)

On hearing the pronunciation [pli:z], the concept we intend for it is “please”. However, when the word is pronounced in a non-prototypical way, either for phonetic or sociolinguistic reason, it is possible that the concept of “police” may come to mind in association with the non-prototypical sound, thereby prompting incongruity and probably humour in certain contexts. This mechanism is found to work in a wide range of humorous instances, especially in interactional humour, in languages across the world. It is not infrequent that the prototypical pronunciation is deliberately deviated to elicit humour. Put differently, the pronunciation of a word is manipulated so that it deviates from the prototypical standard pronunciation and brings about extra funny associations. For instance, “shào nián” means ‘teenager’ or more broadly ‘young people’ in Chinese, but on the Internet it is often deliberately pronounced as “sāo nián” by removing the retroflex consonant, in which “sāo” as an adjective evokes the following connotations: 1) restless or turbulent; 2) coquettish. It is targeted at youngsters who are apparently restrained but actually fiery and whimsical and would behave quite unexpectedly on some occasions. This expression is used to mock their foolish and hilarious behaviour. Chinese as a tonic language is well suited for this strategy. Nowadays, it is common that abundant coined expressions of this kind prevail on the Internet.

Another case of humour built on phonological prototypicality versus non-prototypicality is homonymy, which is the relation between two words that are spelled or pronounced in the same way but differ in meaning. There are perfect homonymous words which share the same spelling and pronunciation (e.g., *bank*) and partial homonymy which share the same pronunciation only (such as *see* and *sea*). In the latter case, the same phonological form might be associated with entirely different meanings/concepts that exhibit typicality more often than not. Compatible with the context and contingent on the communicators, these concepts will compete for representation in the communicators’ minds. If the one that stands out does not fit the expected one, humour is likely to arise. Partial homonymy can be utilised to produce humour by activating the unexpected meaning for the spelling or pronunciation, as Figure 5 shows:



Figure 5. *xìng fú* and *mǎn zú*. Source: Mohun Animation

In this example, the pronunciation of *xìng fú* in Chinese has at least three morphological candidates which share exactly the same pronunciation but have widely distinct meanings (in the sequence of typicality based on frequency): *xìng fú*<sup>1</sup> can function as a noun (meaning happiness) and an adjective (meaning happy or being happy); *xìng fú*<sup>2</sup> likewise functions both as a noun and an adjective, meaning “sexual satisfaction” or “sexually satisfied” respectively;

*xìng fú*<sup>3</sup> is a noun which means “Fu as a surname”. The pronunciation of *mǎn zú* likewise has two meanings: satisfied or content (*mǎn zú*<sup>1</sup>) and Mǎn ethnic minority (*mǎn zú*<sup>2</sup>). The interviewee (with the full name “Erkang Fu”) is a well-known character in a Chinese TV drama, who is cast as a *mǎn zú*<sup>2</sup> nobility in the court of the Qing Dynasty. The journalist’s intended meaning for *xìng fú* and *mǎn zú* is the typical one while what the interviewee takes, from his own background, is the less prototypical one for each. In this text-image humour, the various concepts associated with *xìng fú* and *mǎn zú* are compatible, which help carry on the conversation. But the incongruity involved eventually brings the conversation to a dead end and produces humour. This interpretation is applicable to homophonic pun.

It is evident from the above analysis that phonological prototypicality versus non-prototypicality will contribute to the humorous effect, by drawing on the association between the phonological form and meaning of a word. It is speculated that this humour device will also work at linguistic units larger than a word, for instance, at the phrasal level.

### 3.1.2. Semantic prototypicality versus non-prototypicality

In this kind of multimodal humour device, prototypicality versus non-prototypicality of semantic categories will be exploited, with polysemy as the major player. The various senses of a polysemous word can be described as a prototypical category: the prototypical meaning takes a central position and constitutes the point of departure for semantic extension which is motivated by cognitive mechanisms such as metaphor, metonymy, image schema transformation, and non-cognitive principles such as instantiation, etc. The result is a radial-structured meaning chain (Taylor 2003: 110) as exemplified in Figure 6.

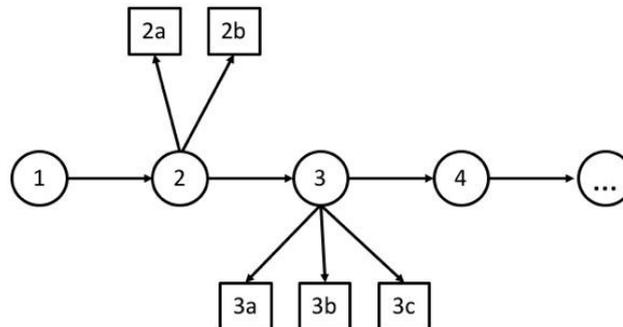


Figure 6. An example of the meaning chain of a polysemy

Normally, we, as speakers, expect to get the prototypical sense across, however, what the listeners capture might be wide of the mark, possibly landing in on any one of the non-prototypical senses that are related to the central one in one way or another. Alternatively, when one of the non-prototypical senses is firmly entrenched in some contexts, but it is the prototypical meaning that is presented, then an incongruity is inevitable. The relations held between the senses render them compatible with the given situation. For instance, in present-day Chinese society, a singleton is jokingly dubbed “a single dog”, in which “dog” refers metaphorically to a person and makes the peripheral member of dog’s semantic category (similar to “poor/lucky dog”). This figurative expression is so established that what conjures up on hearing or seeing the phrase is the image of a single person. However, people will find the captioned picture in Figure 7 humorous because what is presupposed in the verbal modality is the personified sense of “dog” whereas the presented “dog” in the picture traces back to its semantic prototype, which is far beyond expectation. An interpretation of this kind also accounts for homographic pun.



Figure 7. Astonished to meet a single dog in the bar. He is a dog with stories. Source: Weibogaoxiaopaihangbang's post on 14 January, 2015 on Sina microblogging (<http://ww2.sinaimg.cn/large/6628711bgw1eo8z3ux72ej20gt0imq4u.jpg>)

We focus on lexical polysemy in illustrating the humour based on semantic prototypicality versus non-prototypicality. Given that polysemy exists not only at the lexical level of a language, it is natural that this humour device is common in communications, as will be demonstrated in the syntactic, or rather constructional level in the following section.

### *3.1.3. Syntactic prototypicality versus non-prototypicality*

Prototype effects have also shown up in syntax (Lakoff 1987: 61). Construction, understood as a pairing of form and meaning and regarded as a gestalt, is proposed as a one-for-all resolution for the interpretation of various language units, from morpheme, word, even basic sentence patterns through to discourse. In our view, the notion of construction makes a powerful tool to account for humour at the syntactic level since it is taken to be the basic unit of syntax (Taylor 2003: 226). Constructions are productive; just like other kinds of linguistic objects, a construction should be viewed as a prototypical category (Taylor 2003: 222). It follows that two of the most central properties of category structure, i.e., polysemy and prototype-extension structure, are present in construction taxonomies (Croft & Cruse 2004: 273).

As for syntactic constructions, they tend to be complex, abstract and (mostly) schematic (Croft 2004: 275) and may differ in size, complexity and productivity. Constructions might be substantive and lexically filled, such as morpheme, word and compound word; they might be schematic as well, with slots to be filled by lexicons, such as the ditransitive construction. Lexically open, schematic constructions normally allow more than one instantiation and are distinguished according to degrees of schematicity by Fillmore et al. (1988). Various instantiations, also constructions by themselves, will form a network which displays the prototype effect, with some instantiations counting as better examples of the construction than others (Taylor 2003: 222). The prototypical instantiation of the construction is deeply rooted in our mind as a paragon that showcases the most typical combination of form and meaning for the schematic construction. The prototype affords the possibility to simulate its pattern in generating other similar, humorous and often impressive instantiation sentences. This is a

strategy frequently adopted in advertisements nowadays. In the formal dimension, a construction can be seen as a formula consisting of an ordered sequence of slots, some of which are obligatory to the construction while others might be optional (Taylor 2003: 225). Accordingly, one can produce humorous sentences either by simulating the prototypical instantiation at large or by replacing the prototypical item of a slot with non-prototypical candidates, which is analogous to the working mechanism of parody. When we come across parodic instances, we are inclined to compare them with the existing prototype, and will probably detect both similarity to and departure from the prototype. An outstanding example is the schematic construction of Shakespeare's "To be or not to be, that is the question", in which the slot occupied by "be" is replaced by other items and the humorous effect is achieved. An illustrative example is provided in the example in Figure 8, which also takes advantage of the phonological similarity between "be" and "pee".



Figure 8. To pee or not to pee, that is the question. Source: <http://imgur.com/Pw7emCV>

In the semantic dimension, former expositions imply that the meaning of a construction comes from triple interactions—the one between the constituents of the construction, the one between the construction and its constituents, and the one between the construction and the linguistic environment and context of situation it is located in. Goldberg (1995: 31) proposes a polysemy link between the senses of a construction and demonstrates with the example of ditransitive construction how a construction can associate with a family of distinct but related senses: a certain sense of the construction turns out to be more central and prototypical while others are derived from the prototype by means of the aforementioned extension mechanisms, in much the same way as a polysemous word acquires its senses. Taylor (2003: 225) adds that the meaning of a construction should embrace both pragmatic and discourse-related matters. Nonetheless, what should be borne in mind is that no dedicated syntactic structures or configurations always trigger humour (Bergen & Binsted 2015:55); it is when the employment of non-prototypical meaning or usage is not anticipated or at odds with the context that humour will probably be produced. For example, pragmatically speaking, an interrogative syntactic construction is typically used to ask for information. Meanwhile, it can be used for exclamation and other non-prototypical situations, say, the much cited "Waiter, what is this fly doing in my soup?". The syntactic pattern of "Does your family know that ... (usually something trivial or funny)?" plus a supporting illustration, which is not intended to ask for

information, has been popular for expressing humour on the Internet for quite a while, as exemplified in Figure 9.



Figure 9. You are so fat. Does your family know that? Source:  
<http://forum.boolean.name/showthread.php?p=277540>

To round up, prototypicality versus non-prototypicality at the phonological, semantic and syntactic levels of the language, which has long been regarded as an essential element in creating the humorous effect, can be directly or indirectly responsible for the incongruity that leads to humour. The links between the prototypical case and the non-prototypical case, i.e., family resemblance, make them compatible in the context of situation. In short, looking at multimodal humour through the lens of prototype theory provides an alternative perspective to view traditional humour studies.

### 3.2. Humour and non-linguistic category

Non-linguistic categories refer generally to natural concrete categories, such as those in Rosch and her colleagues' study (Rosch et al. 1976), and the abstract categories. They are denoted by linguistic expressions. Firstly, along the vertical axis in Figure 1, other things being equal, people would usually prefer basic-level terms for describing the world, since the superordinate level is too general or abstract whereas the subordinate level is too detailed or specific. In other words, in our daily communication, with three levels of descriptive devices available, one of the communicators might by default choose to talk at the basic level. However, if the other communicator insists on the superordinate or subordinate level, humour would ensue as such. Typical examples, although involving only the verbal modality, can be found in the American sitcom *The Big Bang Theory*, where the audience are sometimes amused by Sheldon's constant use of academic terminologies, most of them being subordinate terms, in his talk to his layman communicator.

Secondly, along the horizontal axis in Figure 1 and in line with the features of the prototype theory, the remainder of this section will delve into how multimodal humour can be built on: 1) prototypicality and non-prototypicality of category members; 2) the family resemblance shared by category members; 3) the fuzzy inter-categorical boundary. The

overarching concept of schema will be invoked for detailed interpretation. Martin (2007: 86) offers an insightful account of the relationship between schema and incongruity:

Schemas describe the general characteristics of an object or event and contain variables or slots that can assume different values in particular instances. For example, a schema for birds would include variables such as the types of wings, feet, beaks, tails, and bodies, which may be instantiated in a number of ways in individual birds. Many different kinds of birds all fit the general schema, with different values for the different variables. The variables often contain default values that represent the prototypical characteristic of the object or event. When we catch a glimpse of a bird or hear about a bird in a story, the schema for birds is activated, and, unless we are given information to the contrary, we expect this particular bird to conform to the default values. The acceptable values of variables in a given schema have certain limits. If we see a drawing of a bird with wings that look like airplane propellers, this would not fit the expected values of the bird schema, and would therefore be an incongruity, something that “does not compute” with respect to our mental model of birds.

(Martin 2007: 86)

Besides, he adds that the schema (script) also specifies the appropriate and relevant details of the narrative in relation to the schema, and the evaluation of people’s actions in it, which enables the narrator to leave out many details that we automatically fill in as defaults (Martin 2007: 86). Speakers in daily communication constantly activate schemas, and a modification of the frame that a speaker maintains is often followed by some humour effect (Bergen & Binsted 2015: 64). The example in Figure 10 speaks well for this point. The schema of having acoustic experience with ears will guide communicators to expect this activity; however, as it unfolds, the anticipated result is subverted, leading to an incongruity in this situation. Readers are forced to reassess the situation and realise that the animals in the cartoon are personified. Similar instances include the popular humorous series on the Internet entitled “ruining my childhood (memory)”, or “mind-boggling series”, which recaption familiar and regular scenarios of cartoons, caricatures and TV dramas in an unexpected way to create humour.

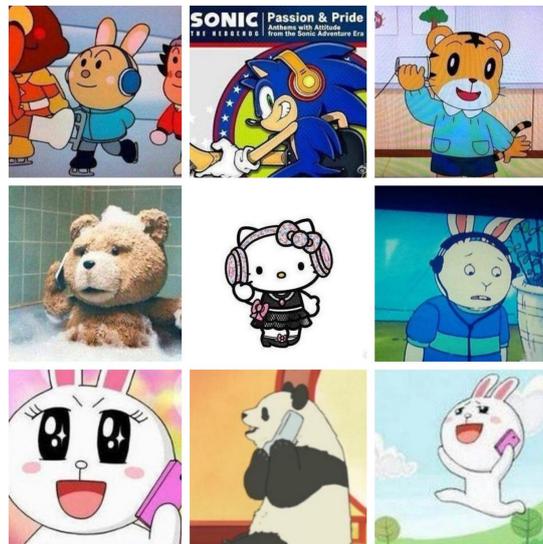


Figure 10. Less educated as I may be, don't fool me, where on earth are your ears? Source:

Yilin Magazine's post on 12 October, 2015 on Sina microblogging

([http://www.weibo.com/yilinzazhi?is\\_all=1&stat\\_date=201510&page=6#\\_rnd1476562219646](http://www.weibo.com/yilinzazhi?is_all=1&stat_date=201510&page=6#_rnd1476562219646))

In what follows, we will examine how the principles of the prototype theory in the case of concrete natural categories could bring about humour, in the sequence of (non)prototypicality, family resemblance and fuzzy boundary.

### *3.2.1. Humour built on prototypicality versus non-prototypicality of category members*

In this case, one of the modalities alludes to a category whose prototype/non-prototype is presupposed while another modality presents the opposite. The incongruity between the expected and the presented, if perceived, will possibly engender humour. In Figure 11, picked out of a humorous series entitled “When breezy words meet ‘unscrupulous’ illustrations”, the context of the caption directs the readers to presume a prototypical member of the “young lady” category. However, the illustration presents a non-prototypical member which turns out to be a man (featuring Chinese actor Huang Bo, from whose Sina posts the picture in Figure 11 is obtained) in women’s dress and contradicts the readers’ expectations.



Figure 11. There will always be a young lady, pure and neat, standing there with a smile at you, the hemline of her shirt blowing with the wind—time stands still at this moment. Source:

Huangbo’s post on 2 August, 2015 on Sina microblogging

(<http://photo.weibo.com/1823463403/wbphotos/large/mid/3871458427980539/pid/6cafd7ebjw1euofyrvi8pj20c80icmzy>)

It is worth mentioning that, in terms of simile and metaphor, the default type of the vehicle is the prototypical member of its own category. For example, in the simile “Her face is like an apple”, what the reader expects of the vehicle “apple” is the prototypical one, i.e., a round-shaped, rosy apple in good condition instead of a rotten or an eaten one. When the illustration

for the verbal simile or metaphor presents a non-prototypical apple which deviates wildly from the prototype, the reader may laugh at the divergence.

### *3.2.2. Humour built on family resemblance*

The family resemblance shared by members of a category can be utilised to create humour. Notwithstanding the family resemblance, category members remain discrete entities. When the distinction is deliberately disregarded and category members are seen as more or less the same, in other words, when the resemblance is highlighted, humour follows from the resultant incongruity. Figure 12 stands out as a good example.



Figure 12. Damn! Handsome again! Source: Yao Chen's post on 29 May, 2011 on Sina (<http://photo.weibo.com/1266321801/wbphotos/large/mid/11398153344/pid/7897b005tw1dhoi5rr0ldj>)

In this example, the kitten and the lion are both members of the cat category and they share some common traits. However, there remain significant differences between them. The reflection in the mirror is expected to be the kitten, and an incongruity arises when we see the image of a lion in the mirror. The resolution of the incongruity lies in that the similarities, or the family resemblance between the kitten and the lion are emphasised over the differences, whereby the mirror creates the illusion for the kitten.

### 3.2.3. Humour built on fuzzy boundaries

The vague boundary between categories is characterised by some peripheral members which possess features of both categories and are often hard to categorise exactly without professional knowledge. These peripheral members, often thought to straddle between the two categories in folk taxonomy, might be wrongly classified. Based on the fact that the boundary between categories are not clear-cut, it is possible that members of one category (in particular non-prototypical members that are closer to the boundary) to count themselves or are counted by others, intentionally or unintentionally, as belonging to either of the adjoining categories. Incorrect categorisation will result in incongruity that leads to humour. A suitable example is given in Figure 13. In China, it is much colder and drier in the north than in the south in winter. However, central heating is generally considered unnecessary in the south as the average temperature and humidity are much higher in winter. People from the north might find the cold and humidity unendurable if they come to the south at this time of the year. In such a context, they would compare themselves to “a wolf from the north”, which is the title of a popular song, and use this picture to express their feelings. It is well known that wolf and dog are closely related as measured by genetic testing and share a common ancestor. In particular, Husky, the dog in Figure 13, bears much resemblance to the wolf. In this example, the boundary between the wolf and the dog category is blurred, that is to say, a wolf can shift to the dog side in some circumstances.



Figure 13. I am a wolf from the north, only to be frozen to be a (poor) dog in the south. Source:

<http://tieba.baidu.com/p/2095467873>

It is worth mentioning that there are both difference and overlapping between humour based on prototypicality/nonprototypicality of category members (Section 3.2.1) and humour based on the fuzzy boundaries (Section 3.2.3). The former may involve incongruity sourced from intra-category or inter-category contrast. To be specific, the contrasted prototypical and non-prototypical cases might be from the same category. The other possibility is that the prototypical case is contrasted with a peripheral one which bears certain resemblance to the prototype yet belongs to another category. However, humour based on the fuzzy boundaries can only involve inter-category contrast, for instance, the dog and the wolf in Figure 13 come from different categories.

In summary, not only (non)prototypicality of category members, but also the family resemblance shared by them and the vague boundaries between categories are capable of generating humour. The examples in this section are from concrete categories, yet the humour device is supposed to work in the case of abstract categories.

### **3.3. Comprehension of multimodal humour built on principles of the prototype theory**

Languages, particularly those which are typologically unrelated, tend to differ substantially in their phonological and syntactic systems. Therefore, it is predictable that the humorous effect of multimodal humour built on phonological or syntactic (non)prototypicality in one language might disappear in the process of translation. As for the semantic system, even though the universality of embodiment tends to suggest similar lexical semantic prototypes, equivalent lexicons in two languages are likely to take different routes of semantic extension and develop different semantic networks. Therefore, semantic (non)prototypicality-based humour may not be universally transmitted, either.

Some non-linguistic categories may exhibit cross-cultural differences in their prototypes in the first place, while others may share the same prototype yet differ in its cultural connotations. In a broader sense, variables or slots of schemas are open to idiosyncratic values in different cultures, for instance, accepted procedures of an event may vary across cultures. As a result, the appropriateness and comprehension of some multimodal humour involving the principles of the prototype theory do not always hold in another culture, which has an influence on cross-cultural communication and transmission. Sociopragmatic factors (e.g. who is the humour initiator, the audience, or the target) also have a role in this type of humour.

All in all, far from being universal, multimodal humour based on prototype theory principles is to a large extent contextually and culturally bound. The successful communication of the humorous effect, especially for cross-culture communication, will require special techniques.

## **4. Intermodal relationship**

This section concerns semiotic relations between participating modalities of multimodal humour. Traditionally, studies on multimodal humour centre around cartoons and other forms of text-image humour, in which the image is generally treated as subordinate to the text. Attardo & Chabanne (1992: 168) claim that the humorous point in cartoons is created by the text-image interaction: it is either “based on the contrast between a perfectly ‘normal’ caption and an incongruous drawing” or the other way round, that is, the caption functions as “some sort of ‘humorous commentary’ on the otherwise nonhumorous drawing”. Taking into account the contributions made by each modality and the location of the punchline, Samson and Huber (2007:14) categorise the text-image relation in cartoons into three types: (a) the text that contains the punchline is more responsible for the humorous effect than the picture which is an illustration of the verbal joke, (b) both modalities collaborate in producing humour in the way

that the picture provides supporting information not contained in the text, but the punchline of the joke remains in the text; (c) the punchline is in the picture which by itself suffices to create the humorous effect. This categorisation is reducible to two points: (i) both the text and image are essential in contributing to humour; one of them is more important as it carries the punchline; (ii) either one of the modalities can single-handedly create humour, the other one being its semiotic equivalent.

Research findings in multimodal discourse analysis will shed some light on intermodal interactions in multimodal humour. Zhang (2009: 26) distinguishes between complementary and non-complementary intermodal relationships in multimodal discourse. On the one hand, when it is insufficient for a single modality to fully express the intended meaning, the second or even third modality will be employed as a complement. According to the status of composite modalities, the complementary relationship is subdivided into reinforcement and non-reinforcement types (Zhang 2009: 26). Modalities in the former are unequal in status, one of them being the main modality for expressing ideational meaning and the rest as secondary one(s), such as the relationship between speech and gesture. Modalities in a non-reinforcement relationship are joined on an equal footing and are indispensable and coordinate closely in the expression of the overall meaning, such as the relationship between sound and image in a video. On the other hand, a non-complementary relationship refers to the one in which both or all modalities coexist even though one or some of them make no extra contribution to the expression of meaning.

Fusing together these approaches, we attempt to model the intermodal relationships between text and image in this study. There are complementary and non-complementary relationships between the modes. In the complementary relationship, both modalities are essential, and they work together to communicate humorous meanings which are beyond the scope of either of them separately. However, if the punchline is located in one of the modalities, the text and image will be unequal in status, with the one that contains the punchline being primary. Thus the modalities engage in a reinforcement relation, such as the examples in Figures 8 and 10. If the punchline exists in both modes, which are equally indispensable for the humorous effect, they are in a non-reinforcement complementary relationship, such as the examples in Figures 5, 7, 11 and 12. The caption in Figure 12, for instance, adds to the humorous effect generated by the monomodality of image. The non-complementary relationship equates to point (ii) previously mentioned. Most of the examples in this study fall into the complementary category.

Table 1. Intermodal relationship between text and image

relation type	specifications		examples in the study
complementary	reinforcement	$m1(p) + m2;$ $m1 + m2 (p)$	Figure 8, 10
	non-reinforcement	$m1 + m2$ $(p1+p2)$	Figure 5, 7, 11, 12
non-complementary	$m1(p) > m2; m2(p) > m1$		Figure 9,13

Notes: m: modality; p: punchline

## 5. Conclusion

To conclude, we have tentatively integrated the prototype theory into the incongruity theory of humour for the analysis of multimodal humour, involving both linguistic and non-linguistic categories. It has been demonstrated with examples that multimodal humour can be accomplished by exploiting the principles of prototype theory, which are prototypicality/non-prototypicality of category members, family resemblance shared by category members and the vague boundary between categories. Detailed operations are analysed in the framework of the conventional incongruity theory of humour. Within linguistic categories, prototypicality versus non-prototypicality at the phonological, semantic and syntactic levels are a potential source of incongruity for humour. In the case of non-linguistic categories, prototypicality/non-prototypicality of category members, the family resemblance and the vague categorical boundary are found to be the mechanisms of multimodal humour. We have also briefly examined the issue of comprehension and translation of multimodal humour based on the principles of prototype theory. Finally, the intermodal relationship between image and text has been investigated and categorised into complementary and non-complementary types. The limitations of this study lie in its being restricted to humour encoded in the dual modalities of text and image. The cases involving three or more modes in varied forms need more research in the future.

## Acknowledgements

The present article is part of the research project “Non-verbal emotional interaction and the effectiveness of foreign language classroom teaching”, supported by the National Social Science Foundation (15BYY082). We would like to thank the two anonymous reviewers for helpful comments and suggestions.

## References

- Attardo, S. (1994). *Linguistic Theories of Humor*. Berlin: Mouton de Gruyter.
- Attardo, S. (2002). ‘Cognitive stylistics of humorous texts’, in Semino, E. & Culpeper, J. (eds.), *Cognitive Stylistics: Language and Cognition in Text Analysis*, Amsterdam: Benjamins, pp. 231–250.
- Attardo, S. & Chabanne, J-C. (1992). ‘Jokes as a text type’. *Humour: International Journal of Humor Research* 5(1-2), pp. 165-176.
- Attardo, S. & Victor, R. (1991). ‘Script theory revis(it)ed: Joke similarity and joke representation model’. *Humour: International Journal of Humour Research* 4 (3/4), pp. 293–347.
- Bergen, B. & Binsted, K. (2015). ‘Embodied grammar and humour’, in Brône, G., Feytaerts, F. & Veale, T. (eds.), *Cognitive Linguistics and Humour Research*, Berlin: Mouton De Gruyter, pp. 49-67.
- Brône, G. & Feytaerts, K. (2003). ‘The cognitive linguistics of incongruity resolution: Marked reference-point structures in humour’. *Reprint No. 205*, Department of Linguistics, University of Leuven.
- Brône, G. & Feytaerts, K. (2004). ‘Assessing the SSTH and GTVH: A view from cognitive linguistics’. *Humour: International Journal of Humour Research* 17(4), pp. 361-372.

- Brône, G., Feyaerts, K. & Veale, T. (2006). 'Introduction: Cognitive linguistic approaches to humour', *Humour: International Journal of Humour Research* 19(3), pp. 203-228.
- Carroll, N. (1996). 'A note on film metaphor'. *Journal of Pragmatics* 26, pp. 809-822.
- Coulson, S. (2001). *Semantic Leaps: Frame-shifting and Conceptual Blending in Meaning Construction*. Cambridge: Cambridge University Press.
- Croft, W. (2004). *Radical Construction Grammar: Syntactic Theory in Typological Perspective*. New York: Oxford University Press.
- Croft, W. & Cruse, D. A. (2004). *Cognitive Linguistics*. New York: Cambridge University Press.
- Deckers, L. (1993). 'On the validity of a weight-judging paradigm for the study of humour'. *Humour: International Journal of Humour Research* 6(1), pp. 43-56.
- Deckers, L. & Buttram, R.T. (1990). 'Humour as a response to incongruities within or between schemata'. *Humour: International Journal of Humour Research* 3(1), pp. 53-64.
- Dirven, R. & Verspoor, M. (1998). *Cognitive Exploration of Language and Linguistics*. Amsterdam: John Benjamins.
- Edwards, J. L. (2001). 'Running in the shadows in campaign 2000'. *American Behavioural Scientist* 44, pp. 2140-2151.
- El Refaie, E. (2003). 'Understanding visual metaphor: the example of newspaper cartoons'. *Visual Communication* 2, pp. 75-95.
- El Refaie, E. (2004). 'Our purebred ethnic compatriots': irony in newspaper journalism'. *Journal of Pragmatics* 37, pp. 781-797.
- Ferro-Luzzi, G. E. (1990). 'Tamil jokes and the polythetic-prototype approach to humour'. *Humour: International Journal of Humour Research* 3(2), pp. 147-158.
- Fillmore, C. J., Kay, P. & O'Connor, M. A. (1988). 'Regularity and idiomaticity in grammatical constructions: The case of let alone'. *Language* 64(3), pp. 501-538.
- Forabosco, G. (1992). 'Cognitive aspects of the humour process: The concept of incongruity'. *Humour: International Journal of Humour Research* 5(1-2), pp. 45-68.
- Fraser H. (2006.) 'Phonological concepts and concept formation: Metatheory, theory and application'. *International Journal of English Studies* 6(2), pp. 55-75.
- Geeraerts, D. (1989). 'Prospects and problems of prototype theory'. *Linguistics* 27, pp. 587-612.
- Geeraerts, D. (1997). *Diachronic Prototype Semantics: A Contribution to Historical Lexicology*. New York: Oxford University Press.
- Goldberg, A. E. (1995). *Constructions: A Construction Grammar Approach to Argument Structure*. Chicago/London: The University of Chicago Press.
- Hull, R., Tosun, S. & Vaid, J. (2017). 'What's so funny? Modelling incongruity in humour production'. *Cognition and Emotion* 31(3), pp. 484-499.
- Kang, B. (2016). 'Metaphor and its humorousness: The case of nominal compounds in German'. *Humour: International Journal of Humour Research* 29(3), pp. 359-380.
- Labov, W. (1972). *A Quantitative Study of Sound Change in Progress*. Philadelphia: U.S. Regional Survey.
- Labov, W., Cohen, P., Robins, C. & Lewis, J. (1968). *A Study of the Non-standard English of Negro and Puerto Rican Speakers in New York City*. Cooperative Research Report 3288, Vols I and II. Philadelphia: U. S. Regional Survey.
- Lakoff, G. (1987). *Women, Fire, and Dangerous Things*. Chicago: University of Chicago Press.
- Mandler, J. M. (1979). 'Categorical and schematic organisation in memory', in Puff, C. R. (ed.), *Memory Organisation and Structure*, New York: Academic Press, pp. 259-299.

- Martin, R. A. (2007). *The Psychology of Humour: An Integrative Approach*. New York: Academic press.
- Martinec, R. & Salway, A. (2005). 'A system for image–text relations in new (and old) media'. *Visual Communication* 4(3), pp. 337-371.
- McGhee, P. E. & Panoutsopoulou, T. (1990). 'The role of cognitive factors in children's metaphor and humour comprehension'. *Humour: International Journal of Humour Research* 3(4), pp. 379–402.
- Nathan, G. S. (1986). 'Phonemes as mental categories'. *Berkeley Linguistics Society (BLS)* 12, pp. 212–223.
- Nathan, G. S. (2007). 'Phonology', in Geeraerts, D. & Cuyckens, H. (eds.), *The Oxford Handbook of Cognitive Linguistics*, New York: Oxford University Press, pp.611-631.
- Nathan, G. S. (2008). *Phonology: A Cognitive Grammar Introduction*. Amsterdam/Philadelphia: John Benjamins Publishing Company.
- Nikolajeva, M. & Scott, C. (2000). 'The dynamics of picturebook communication'. *Children's Literature in Education* 31 (4), pp. 225–239.
- Norrick, N. R. (1986). 'A frame-theoretical analysis of verbal humour'. *Semiotica* 60, pp. 225-245.
- Plumb, S. (2004). 'Politicians as superheroes: the subversion of political authority using a pop cultural icon in the cartoons of Steve Bell'. *Media, Culture and Society* 26, pp. 432–439.
- Ritchie, Graeme. (2004). *The Linguistic Analysis of Jokes*. London/New York: Routledge.
- Rosch, E. (1978). 'Principles of categorisation', in Rosch, E. & Lloyd, B. B. (eds.), *Cognition and Categorisation*, Hillsdale, NJ: Lawrence Erlbaum, pp. 27-48.
- Rosch, E. & Mervis, C. B. (1975). 'Family resemblances: Studies in the internal structure of categories'. *Cognitive Psychology* 7, pp. 573–605.
- Rosch, E., Mervis, C. B., Gray, W. D., Johnson, D. M. & Boyes-Braem, P. (1976). 'Basic objects in natural categories'. *Cognitive Psychology* 8(3), pp. 382–439.
- Ruch, W. (2008). 'Psychology of humour', in Raskin, V (ed.), *The Primer of Humour Research*, Belin: Mouton de Gruyter, pp. 17-100.
- Samson, A. C. & Huber, O. (2007). 'The interaction of cartoonist's gender and formal features of cartoons'. *Humour: International Journal of Humour Research* 20, pp. 1–25.
- Suls, J. M. (1972). 'A two-stage model for the appreciation of jokes and cartoons', in Goldstein, J. H. & McGhee, P. E. (eds.), *Psychology of Humour*, New York: Academic Press, pp. 81–100.
- Tabacaru, S. & Feyaerts, K. (2016). 'The power of metonymy in humour: stretching contiguous relations across different layers of meaning'. *European Journal of Humour Research* 4(4), pp. 1-18.
- Taylor, J. R. (2003). *Linguistic Categorisation (Third edition)*. Oxford: Oxford University Press.
- Templin, C. (1999). 'Hillary Clinton as threat to gender norms: cartoon images of the First Lady'. *Journal of Communication Inquiry* 23, pp. 20–36.
- Tsakona, V. (2009). 'Language and image interaction in cartoons: Towards a multimodal theory of humour'. *Journal of Pragmatics* 41 (6), pp. 1171-1188.
- Ungerer, F. & Hans-Jörg, S. (1996). *An Introduction to Cognitive Linguistics*. London: Addison Welsey.
- Raskin, V. (1985). *Semantic Mechanisms of Humour*. Dordrecht: D. Reidel.
- Wang, Y. (2006). *Cognitive Linguistics*. Shanghai: Shanghai Foreign Language Education Press.
- Wyer, R. S. & Collins, J.E. (1992). 'A theory of humour elicitation'. *Psychological Review* 99(4), pp. 663-688.

Zhang, D. (2009). 'On a synthetic theoretical framework for multimodal discourse analysis'.  
*Foreign Languages in China* 6(1), pp. 24-30.