The Internal Formation Mechanism of Verbal Humor Level

JIANG Zhen-ting
Hunan University of Humanities Science and Technology, Loudi, China

Humor is a kind of verbal wisdom. In the process of making humor, both the speaker and the listener can feel the difference of humor effect. Through the investigation and research of many humor examples, discussion is made for the internal formation mechanism of humor level from perspectives of both coding and decoding, combined with prototype theory and spatial synthesis theory. The results show that the internal formation mechanism of humor level is mainly implemented in two aspects: first, the distance between the edge association and the prototype association forms a humor level difference, and second, the difficulty level of spatial synthesis forms a humor level difference.

Keywords: humor level, formation mechanism, prototype theory, spatial synthesis theory

Statistical Analysis on the Level Difference of Verbal Humor

In order to deeply understand the level difference of verbal humor, a hundred various humor cases are randomly selected in major newspapers, magazines and on net such as “Humor Master”, “Reader”, “Humor and Joke”, “Stories” and so on. Then, a questionnaire survey on verbal humor was conducted among 200 undergraduate students of different majors. The questionnaire consists of closed structural questions and open non-structural questions. For the question “Do you think the ridiculousness of humor is graded?”, 96% of the students answered “Yes”. Six factors, (1): “target”, such as strong groups, general groups, vulnerable groups; (2): “context”; (3): “humorous methods”, such as satire, ridicule, teasing, teasing, etc.; (4): “wisdom”, which is the imagination of humor and the reader’s mental intelligence to interpret this humor; (5): “subject matter”, such as politics, culture, and entertainment; (6): “the cultural background of the reader”, are sorted by influence to
determine the humor level, and the factors appearing in top three are “wisdom” (136 times), “subject matter” (117 times), and “context” (102 times). The 100 humor cases in the questionnaire are divided into 5 levels according to the level of wisdom, A low (0-2 points), B a little low (2-4 points), C medium (4-6 points), D a little high (6-8 points), E high (8-10 points). When scoring, the statistics show that 11 cases are with an average score of 0-2 points, 28 cases are with an average score of 2-4 points, 42 cases are with an average score of 4-6 points, 14 cases are with an average score of 6-8 points, and 5 cases are with an average score of 8-10 points. Accordingly, the number distribution of the 100 cases in the five wisdom intervals can be displayed in the statistical analysis chart as follows.

![Figure 1. The intelligence distribution table of humorous utterances.](image)

The statistical analysis shows that although there are many factors to form humor level, the internal wisdom is the dominant factor. When constructing verbal humor, verbal humor below medium wisdom is relatively easy to form. Higher wisdom requires broader, deeper and more novel imagination, which leads to more difficulty in formation. Therefore, there are fewer high-level verbal humor cases.

So, how does wisdom form the level of verbal humor? Based on the survey data above, the following views are proposed by combining prototype theory and space synthesis theory.

**The Distance between the Two Edge Associations and the Prototype Associations Forms a Humorous Differential**

When coders write humorous words, the propositions presented must have a relationship with the contextual assumptions, which are called “associations” in relevance theory. Relevance theory believes that the greatest contextual effect with minimal processing effort is the “maximum association”. According to category theory, we can refer to “maximum association” as “prototype association”; because the characteristics of this association are the most vivid and are good samples in all associations. In communicative activities, the effort to obtain “prototype association” is minimal, so this association is the most easily associated conception of human cognition. In general, decoders always speculate the idea of coders with “maximum association” in the context of the proposition. If the information given by coders is in the imagination and inference of decoders, then the code compiled by coders is the prototype association formed by decoders according to the context hypothesis. Take the following cases as an example:

**Example (1)** Two students with very poor grades came together after the exam.

“Jack, how are you doing?”

“Nothing, I made a white roll, you, Siri?”
“Oh, me too!”
“What shall we do? People will say that we are stupid.”

In example (1), under the laying of previous context, the final response is in decoders’ conception. Decoders do not need to use any mind to know the message that “people will say we are stupid.” Amount of information given by the last few words is extremely low. The lower the amount of information is, the poorer the context effect is. This kind of coding method has low humor value if used in humor cases. But if we replace the last sentence with “what shall we do? People will say that we cheated.” This response goes beyond decoders’ assumptions and inferences, and is a dissociation from the above context. However, it is this dissociation that frees it from the hypothetical context and thus carries a new amount of information. These new amounts of information are related to the context in a deeper way, requiring decoders’ reasoning and interpreting; and decoders will be delighted in the process of comprehension, and the sense of humor is thus given birth. The sentence after the replacement was scored at 4.8 in the survey and the humor level was medium.

According to prototype theory, if the final information given by coders is not a prototype association in cognition, but a variety of deviations from the prototype association, these deviations take the prototype as a center and spread as “continuum” to the periphery from the prototype, some edge members even merge into other categories. The farther away from the prototype association, the longer decoders reason and the harder the effort will be; however, the more information they receive, the stronger their sense of self-identification is and the more the return of humor they have. Therefore, the more edged the association is, the further the distance it is with the prototype, the stronger the sense of humor will be. Relevance theory refers to the associations that achieve the above communicative effects as the best association. In association theory, only two forms of association are mentioned, namely the maximum association and the best association. Because of the singularity and generalization of the reference, in pragmatic research, language researchers collectively refer to the association of good pragmatic effects as the best association. Between these two kinds of references, there are many language vacancies; because, since there is the best association, there should be “good association”, “better association” and “much better association” in transition. In speech activities, it is these forms of association that form the humor level.

In order to explain this point better, four cases are selected in survey, which are in the grades of low, a little low, a little high and high. They are explained with the combination of the figures.


(3) An old couple went to take pictures, the photographer asked: “Uncle, do you want sidelight, backlight, or all-light?” The old man said shyly: “I don’t care, but can your aunt keep her pants?”

(4) A: “My parents are strict with me. Every time I have bad scores, I will be beaten at home.”
B: “What happened last night?”
A: “Beginning is a man’s singles, then a woman’s singles, and finally a mixed doubles for man and woman.”

(5) An American who yelled that Bush was a fool in front of the White House was immediately arrested. The crime was - leaking state secrets.

These four humor cases can be analyzed with the figures.
First, case (2) averages 1.3 points in the survey. What the bunny, the piggy, and the chicken say has created a context. Under this context, the associated idea that is easily formed by decoders is: the puppy will say, “I am son of a bitch”. This is the maximum association that can be learned from the previous context, and is indicated by point A in Fig. 2. However, in the context of Chinese culture, this sentence has an extended meaning and is a curse; therefore, the puppy said something to avoid, “You chat, I am leaving.” This sentence has a distance from the maximum association “I am son of a bitch”, according to its score in the survey, it’s marked as point B in Fig. 2. Although there is a distance between point B and point A, as long as there is a relevant cultural background, decoders can push point B to point A with almost no effort. Therefore, the amount of information that humor brings to decoders is almost zero, thus its humor value is not high. In the eyes of rigorous judges, it is not a humor, which is why it scores not high in the survey.

![Figure 2. The distance graph of humor case (2) and maximun relevance.](image)

Case (3) has an average score of 3.1 points. When decoders hear the “light” that the photographer says, the maximum association produced is the meaning of “optical line”, which is still indicated by point A in the figure. But as a multi-meaning word, in addition to the meaning of “optical line”, there is also a meaning of “uncovered”, and the uncle’s answer is obviously using the meaning of “uncovered” to interpret the photographer’s words. Thus, in addition to the prototype association of light, a deviation from the prototype association is created. According to its score in the survey, it is indicated by point C in Fig. 3. But whether it is the meaning of “optical line” or the meaning of “uncovered”, it is also taken into account in the meanings of light; therefore, its degree of deviation is not high, and listeners can understand the meaning with a little thought. However, after all, it has a distance from the prototype, which deviates from the usual thinking mode of listeners, thus brings the difficulty of interpretation, prolongs the time of interpretation, and the rewards obtained are of course higher. Therefore, the humor value is obviously higher than case (2).

![Figure 3. The distance graph of humor case (3) and maximun relevance.](image)
Case (4) has an average score of 6.9 points. The humor method in case (4) is the same as in case (3) in using the multiple meanings of words to deviate from the prototype association; but when it creates a new association, in addition to deviating from the prototype association, another language domain has also been introduced, which is from hitting people into ball sports. Based on its score in the survey, the prototype association is indicated by point A in Fig. 4, and the edge association created in the case is indicated by point D. As can be seen from Fig. 4, due to the semantic offset and the span of language domain, D is farther away from the prototype. The farther the distance is, the more mental effort decoders spend, the longer it takes, the more intensity of wisdom is needed, and the greater the pride of self-perception will be. The humor value of this cases is obviously higher than in case (2) and (3).

Case (5) has an averages of 9.2 points in the survey. “An American who yelled that Bush was a fool in front of the White House was immediately arrested.” This is the context the case gives us. Under this context, it activates the illegal behavior of the man. As for what kind of illegal behavior, the prototype association formed by decoders according to the context is defamation. But the crime is actually leaking state secrets. Leaking state secrets is of course far from the prototype association, but how far is the focus of the next analysis, and the reason for the highest humor value. It can be seen from Fig. 5 that although the crime of defamation and revealing state secrets are separated by a certain distance, they are still in the same language domain, and the distance between them is not more than the distance formed by the two associations in case (4). But in the first three cases, no matter how far the two related distances are, their cognitive focus is always consistent, and they highlight the narrative subject of discourse; while the cognitive reference point created by case (5) has undergone fundamental changes. In the previous context of case (5), “an American” is the conceptual focus of decoders’ cognition. The American activity, yelling, activates another factor in the context, the president. Next, coders change the cognitive focus quietly, and the “president” became the focus of cognition. If the first two cases are linear in the formation of a prototype, it is piecewise linear here. The quiet transformation of semantic orientation has greatly broken through the cognitive assumptions of listeners. The initial judgment of listeners is suddenly turned over, stepped on the air, and there was no reliance on it. Cognitively, there was a sense of incomprehension. Decoders need to spend more time to think about the intent of coders. After tortuous thinking to get the verbal intent of coders, his sense of pleasure will break forth like an oil well. What’s more, it is the president who is high above the social class that is ridiculed. To ridicule a man who is much stronger than himself, coders and decoders can
gain spiritual sense of victory in addition to mental satisfaction. This is the reason why satirical works have a stronger sense of humor.

![Figure 5. The distance graph of humor case (5) and maximum relevance.](image)

In the above four humor cases, the formation mechanism of humor level is analyzed from the perspective of coders. Through analysis, it is found that in the process of coding, if the association created by coders is farther away from the prototype according to the context, the level of verbal humor is higher, which shows the proportional relationship of them.

**The Difficulty of Spatial Synthesis Forms Humor Level Gradient**

Wang Yan thoroughly explained the contextual hypothesis process of humorous language in “The Pragmatic Essence of Humorous Speech and Contextual Hypothesis Flow” (Wang, 2005). She believes that the realization of humor depends on the audience, and the audience’s perception of humor depends on perception. It is true that the generation of humorous speech requires three presuppositions: coders, decoders, and the context. In the same sentence, the speaker may not feel humor, the listener feels; or the speaker deliberately creates humor, the listener is indifferent; or the speaker thinks this kind of humor, the listener thinks that one and so on. Without the participation of decoders, wonderful humor can only get the effect of asking way to the blinds. Therefore, the formation of humor level is closely related to the psychological cognition of decoders. From the perspective of decoders, there are two situations in which the humor level is formed: one is that same humor forms humor level gradient due to difference in the cognitive level of decoders, which is not discussed here; the other is that different humor forms humor level gradient to the same decoder because of different psychological space to be formed. This is what will be searched deeply here.

In order to explore more clearly the psychological mechanism of humor level formation from the perspective of decoders, we introduce the “synthesis space theory” (Fauconnier, 1997). Synthetic space theory is created by famous cognitive linguists Faunnier, Turner and others based on the theory of psychological space. It is considered a basic and universal cognitive way of human beings. The psychological space in cognitive linguistics refers to the Conceptual Packet constructed by people for the purpose of local understanding and action when thinking and talking. Psychological space can establish a series of concepts, such as time, belief, desire, possibility, virtuality, location, reality and so on. People’s daily communication is the result of organic connection by the psychological space words are constantly activated. Synthetic space is the psychological space constructed by human thinking to map and integrate, and to create a conceptual meaning space containing a new organizational structure. According to Faunnier and Turner’s hypothesis, synthetic space consists of a series of
mental spaces (generic space, input space I, input space II and synthetic space). Each mental space consists of partial assemblies of various elements, consisting of framed and cognitive patterns, which are interconnected to form a network, modified with the mind and text, as shown in Fig. 6:

![Figure 6. Synthetic space composition diagram.](image)

In real life, due to the diversity and complexity of human cognition, input space may not necessarily consist of two, and can be mapped simultaneously by multiple input spaces.

The theory shows that spatial synthesis theory is the key to the black box of cognitive psychology of decoders. Proper use of the key can explore the internal mechanism of verbal humor level formation from the perspective of decoders.

The analysis of cognitive psychology of decoders is still carried out by taking the above four cases as an examples.


The four psychological spaces are shown in Fig. 7. The structure in generic space is that “the fosterer has raised her own children”. Input space I represents the statement four animals make in the case, and input space II is the cognitive domain of people. The abstract structure in generic space determines the one-to-one correspondence between input space I and input space II.
As shown in Fig. 7: According to the case, the bunny, the piggy, the chicken respectively state their own fosterers. In the process of understanding, decoders naturally construct the generic space of “foster” and inputs A1, b1, c1 in space I and a2, b2, and c2 in space II. According to the statements of the first three small animals, their claims are consistent with the cognitive processes in people’s minds. Therefore, the connection is established in synthetic space, and the first step of integration process is completed: combination. The synthesis of a1a2, b1b2, and c1c2 is very clear according to the connection of the figure, and does not require any mentality. Next, the puppy’s words, d1 also entered input space I. At this point, even if the puppy does not say, according to previous projection principle, it is natural to form a synthetic conclusion of d1d2 in synthetic space. But the puppy apparently doesn’t say the synthetic conclusion of d1d2, but “You chat, I am leaving” instead. Thus, the puppy’s words are obviously different from d1d2. At this time, decoders must start other knowledge to find the reason for the difference between d1d2 and the puppy’s words. “son of a rabbit”, “son of a sow” and “son of a hen” have no extended meaning in daily communication; however, “son of a bitch” has a certain extended meaning, that is, d1d2 composite line is a two-wire system with a dark line below the open line. To discover the dark line needs to mobilize cultural background knowledge to see that “son of a bitch” is gross. And that’s why the puppy doesn’t say that. Through the analysis, it is found that to understand this humor, the mentality decoders spend is mainly on the double-line search for d1d2, and little on other places. The clues already exist, when using existing background to find, it is relatively easier than re-establish other connections; therefore, its humor level is low.

Example (3) An old couple went to take pictures, the photographer asked: “Uncle, do you want sidelight, backlight, or all-light?” The old man said shyly: “I don’t care, but can your aunt keep her pants?”
The four psychological spaces are shown in Fig. 8. The structure in generic space is “a person adopts a certain way when doing something”, so the shared organizational framework is “a person does something in a certain way.” Input space I is the cognitive domain of the photographer, input space II is the cognitive domain of the old couple, and the abstract structure in generic space determines the one-to-one correspondence between input space I and input space II.

As shown in Fig. 8, “photographer” and “the old couple” are participants in the activity, and correspond to each other, “photographing light mode” corresponds to “how much clothing at the time of shooting”. In the process of understanding, the first step of integration process is completed: combination. Anyone who has a certain understanding of photography knows that the photographer’s “sidelight, backlight, or all-light?” is a term used in photography. Common background knowledge triggers the readers’ common cognitive system and completes the second step of integration: perfect. In synthetic space, photographer and the old couple have achieved unity in their behaviors. However, due to different knowledge backgrounds, the old couple is different from the photographer in understanding the same thing. Different understandings lead to inability of c1c2 to effectively map in two spaces. Therefore, in synthesis space, although a1a2 and b1b2 can be integrated, c1c2 is still separated like oil and water (this relationship is indicated by a dotted line in Fig. 8). This requires decoders to construct a new meaning association, that is, to effectively link c1c2 by understanding the lighting method as how much to dress, thus to complete a new definition of light. The formation of new meanings means that the conceptual integration has to complete the final step: development. Compared re-establishing associations with finding associations, the former consumes more mentally. Usually, the humor of speech, the wisdom of input by coders and the wisdom of decoders’ interpretation are directly proportional; therefore, the humor level of this humor is obviously higher than that of case (2).

Example (4) A: “My parents are strict with me. Every time I have bad scores, I will be beaten at home.”
B: “What happened last night?”
A: “Beginning is a man’s singles, then a woman’s singles, and finally a mixed doubles for man and woman.”

The four psychological spaces are shown in Fig. 9. The structure in generic space is “punished by parents for some reason”, and the shared organizational framework is “punished by parents due to unsatisfactory test scores”. Input space I is the cognitive domain of B, and input space II is the cognitive domain of A. Literally, input space I is of course a kind of space that can enter the parental education of children, but input space II has no connection with the education of children (the relationship is indicated by the dotted line in Fig. 9). In this way, decoders need to spend the mind to establish their intrinsic connections. The abstract structure in space determines the one-to-one correspondence between input space I and input space II. However, after a1 and a2 enter the synthesis space, synthesis depends on the decoders’ connection to input space II and the generic space. If the association is not enough, the integration will be weakened and the perception of humor will be reduced. It can be seen from the above analysis that this humor brings difficulty to decoders in two places, and the mentality that decoders has to successfully decode it is obviously higher than in case (3), and the pleasure generated for this is of course larger than case (3).

Example (5) An American who yelled that Bush was a fool in front of the White House was immediately arrested. The crime was - leaking state secrets.

The four psychological spaces are shown in Fig. 10. The structure in generic space is “someone does something to someone and therefore gets the corresponding result.” The shared organizational framework is “someone has cursed someone.” Input space I is the cognitive domain of general citizen, input space II is the cognitive domain of this American, input space 3 is the cognitive domain of American citizens, and the abstract structure in generic space determines that input space I correspond one-to-one to input space II.
As shown in Fig. 10, a1a2a3, b1b2 are all compatible with each other, but b1b2 and b3 cannot correspond, and decoders need to be mentally connected (connected by a dotted line in Fig. 10). Next, c1c2 can correspond, but c1c2 and c3 obviously cannot correspond, and decoders need to spend the mind to connect the three (this relationship is connected by a dotted line in Fig. 10). According to Fig. 10, d1d2d3 can be corresponding, and decoders do not need to consume mentality, but whether e1e2 and e3 can be effectively linked depends on how the previous b1b2 and b3 and c1c and c3 are linked. If the previous link cannot be effectively linked, there is a problem with the integration of e1e2 and e3 (this relationship is connected by a dotted line in Fig. 10). Only when all the relationships are effectively linked, all the elements of the three input spaces can be effectively integrated to form a new concept space. In synthetic space, “citizens” and “Americans” act as senders of the behavior, and the results of “arrest” can be unified. But whether “curse” and “leak state secrets” can be integrated must be based on the previous effective association (the uncertainty of synthesis is connected by a dotted line in Fig. 10). “Citizens” and “Presidents of the US” have internal and external unity as the object of behavior, but the two are not uniform with the “state secrets of the US president is a fool”. The internal unity requires decoders to spend their mind to established (the uncertainty of synthesis is connected by dotted lines in Fig. 10). In input space I and the input space II, “defamation” as the reason for arrest, is a fusion from the external form, thus can naturally be effectively integrated in the new layer creation space. However, “leak state secret” in input space III has a clear distance from the other two. To achieve successful integration, decoders need to spend their mind (the uncertainty of the synthesis is connected by dotted lines in the figure). In case (5), there is a lot of uncertainty in the spatial combination of humor in the process of decoding. The connection of many nodes requires mentality; therefore, decoders obtain pleasure and accomplishment after successful decoding, and its humor level is even higher.
From the perspective of decoders, spatial synthesis theory is used to explore the internal formation mechanism of verbal humor level. Through the analysis of four cases, it is found that in the process of perceiving humor, the more nodes decoders need to connect to form a new synthetic space, the stronger the pleasure obtained after successful decoding, and the higher the value of verbal humor. It can be seen that the difficulty gradient of spatial synthesis is another important factor in the formation of humor level.

Conclusion

From the above analysis, we can see that the level of verbal humor is closely related to the psychological creation of the coders and the psychological process of the decoders. On the one hand, the greater the distance between the relevance created by the encoder and the prototype of the decoder, the higher the verbal humor index. On the other hand, in the process of perceiving verbal humor, the more nodes the decoder needs to connect to construct a new synthetic space, the higher the pleasure of decoding and the value of verbal humor.

References


