Animacy hierarchy within inanimate nouns: English corpus evidence from a prototypical perspective

Jie Ji (China Foreign Affairs University, China) and Maocheng Ling (Beijing Foreign Studies University, China)

Animacy has usually been regarded by linguists as a gradient hierarchy rather than a simple biological dichotomy. The animacy of an entity depends not on its biological attributes, but on "the speaker's identification or empathy" with this entity (Kuno and Kaburaki, 1977: 628), or "to what extent speaker treat this entity as if it was animate" (Rosenbach, 2008: 154). Animacy hierarchy can be seen as the gradient assessment towards all entities from the speaker's point of view, and it proves to be widely existent in human languages and highly influential to many grammatical phenomena. Typological studies find that animacy hierarchy has exerted considerable constraints on phenomena like case, number, agreement, voice and syntactic order (Aissen, 2003; Comrie, 1989; Corbett, 2000; Jelinek and Demers, 1983; Oshima, 2007), and these constraints prevail in human languages as either grammatical rules or linguistic preferences (Bianchi, 2006; Bresnan et al., 2001; Rosenbach, 2008).

Drawn from those linguistic manifestations or cognitive assessments, several animacy hierarchies have been proposed (Corbett, 2000; Dixon, 1979; Langacker, 1991; Silverstein, 1976; Yamamoto, 1999). These hierarchies are fine-grained only within the animate entities. Person pronouns, proper nouns, human-referring nouns and animals are granted different ranks along the scale, leaving the inanimate entities "an undifferentiated class" with only few "arbitrary" distinctions (Comrie, 1989: 197). Does this mean that all inanimate entities are equally inanimate without any hierarchic nature? Or that their hierarchy is simply not worth exploring for its insignificant impacts on human languages? These are the two questions the present research is going to investigate.

According to the cognitive approach (Dahl, 2008; Langacker, 1991; Yamamoto, 1999), animacy hierarchy reflects human's egocentric or prototypical assessment of entities. Animacy of an entity depends on how much resemblance it bears to us, or more specifically, how many human prototype features it possesses. The animacy of inanimate nouns could be gradient if they are assigned different amounts of human features by us. Human is agentive, definite, individual, kinetic, cognitive and social (Aissen, 2003: 437; Dahl, 2008; Dixon, 1979: 85; Gelman et al., 1995; Johnson et al., 1998; Woodward, 1998). However, these simple lists of prototype features are inadequate to predict a credible hierarchy, as the amount and strength of each category's features are difficult to measure in this way. We need a feasible method to externalize and quantify these prototype features.

When we say humans are physically kinetic or psychologically cognitive, sentient, communicative and social, we mean they can walk, think, love, exchange ideas and socialize with each other. In other words, humans are "teleologically capable of generating these actions" (Folli and Harley, 2008: 192), that is, humans have the inherent attributes to participate in the actions described by these predicative verbs. In this sense, human attributes or the prototype features can be externalized linguistically as human subjects' predicative verbs. Whenever a human subject occurs with a predicative verb, we can add that verb to the feature list. Then the amount of prototype features can be easily quantified as the total occurrences of these verbs.

This method can calculate not only the amount, but also the strength of prototype features. For example, humans can think and change, thus the quality to "think" and

"change" can both be regarded as human attributes or prototype features. However, the two features have different weights or "cue validities" in their contribution to animacy. Just as "gills" is more valid than "oviparity" when labelling the fish category (Lakoff, 1987: 53), the ability to "think" is much more valid than the ability to "change" when identifying human category. The validity of each cue depends on how frequently "it is associated with members of that category" (Evans and Green, 2006: 261). Thus, the weight of each prototype feature can be calculated by how often its externalized predicate co-occurs with human subjects. More weighting should be given to those more valid features when calculating the amount and strength of each noun category's prototype features.

By using predicative verbs as linguistic externalization of prototype features, this research intends to investigate the animacy hierarchy within inanimate nouns through human language. English is used as the experimental language in this study. The corpus data come from the written part of British National Corpus (BNC). WordNet is used as a noun database in this study. It classifies all nouns into 25 groups, and lists each sense of each noun with group label and corpus frequency. This is useful for our noun grouping procedure. Stanford Parser is used to parse grammatical relations for corpus data. It can provide dependencies such as "nsubj", which labels the syntactic subject and predicate in a clause. We can use this dependency to calculate how often each predicative verb co-occurs with human subjects (hence the weight of each feature), or how frequently and strongly each inanimate noun category co-occurs with these verbs (hence the animacy of each category). We also develop an animacy calculating software that can search through the parsed corpus and automatically calculate the occurrence, weight and animacy, then finally provide a Microsoft Excel output according to our design.

The results attest to the very existence of the animacy hierarchy of five inanimate categories. Collective nouns, spatial and temporal nouns, concrete nouns, psychological nouns, and other abstract nouns rank in descending sequence along the hierarchy. This hierarchy is determined directly by each category's feature distribution in different weight intervals, and the semantic distribution helps to understand the underlying cognitive mechanism of this hierarchy. Further syntactic exploration show that this hierarchy can impact on word order just as those existing hierarchies do.

Reference

Aissen, J. (2003). Differential object marking: Iconicity vs. economy. *Natural Language and Linguistic Theory, 21 (3),* 435-483.

Bianchi, V. (2006). On the syntax of personal arguments. *Lingua, 116 (12),* 2023-2067.

Bresnan, J., Dingare, S., & Manning, C. D. (2001). Soft constraints mirror hard constraints: voice and person in English and Lummi. In: Butt, M., & King, T. H. (Eds.), *Proceedings of the LFG 01 Conference* (pp. 13-32). Stanford: CSLI.

Comrie, B. (1989). *Language Universals and Linguistic Typology, 2nd edition*. Chicago: University of Chicago Press.

Corbett, G. G. (2000). *Number*. Cambridge: Cambridge University Press.

Dahl, Ö. (2008). Animacy and egophoricity: Grammar, ontology and phylogeny. *Lingua, 118 (2)*, 141-150.

Dixon, R. M. W. (1979). Ergativity. *Language, 55 (1)*, 59-138.

Evans, V., & Green, M. (2006). *Cognitive Linguistics: an introduction*. Edinburgh: Edinburgh University Press.

Folli, R., & Harley, H. (2008). Teleology and animacy in external arguments. *Lingua, 118 (2),* 190-202.

Gelman, R., Durgin, F., & Kaufman, L. (1995). Distinguishing between animates and inanimates: Not by motion alone. In Sperber, D., Premack, D., & Premack, A. J. (Eds.), *Causal cognition: A multidisciplinary debate* (pp. 150-184). Oxford: Clarendon.

Jelinek, E., & Demers, R. (1983). The agency hierarchy and voice in some Coast Salish languages. *International Journal of American Linguistics, 49 (2)*, 167-185.

Johnson, S., Slaughter, V., & Carey, S. (1998). Whose gaze will infants follow? The elicitation of gaze-following in 12-month-olds. *Developmental Science, 1 (2),* 233-238.

Kuno, S., & Kaburaki, E. (1977). Empathy and syntax. *Linguistic Inquiry, 8 (4),* 627-672.

Lakoff, G. (1987). *Women, Fire and Dangerous Things: what categories reveal about the mind*. Chicago: University of Chicago Press.

Langacker, R. W. (1991). *Foundations of Cognitive Grammar, Vol. 2: Descriptive Application*. Stanford: Stanford University Press.

Oshima, D. Y. (2007). Syntactic direction and obviation as empathy-based phenomena: a typological approach. *Linguistics, 45 (4)*, 727-763.

Rosenbach, A. (2008). Animacy and grammatical variation -- Findings from English genitive variation. *Lingua, 118 (2)*, 151-71.

Silverstein, M. (1976). Hierarchy of features and ergativity. In: Dixon, R. M. W. (Ed.), *Grammatical Categories in Australian Languages* (pp. 112-171). Canberra: Australian Institute of Aboriginal Studies.

Woodward, A. (1998). Infants selectively encode the goal object of an actor's reach. *Cognition, 69 (1),* 1-34.

Yamamoto, M. (1999). *Animacy and Reference: A cognitive approach to corpus linguistics*. Amsterdam: John Benjamins.