Gradience, features and hierarchies*

Ida Toivonen

Prepublished version
Final version in *Theoretical Linguistics* 44(1–2): 93–98

1 Introduction

In *Shifting Animacy*, Peter de Swart and Helen de Hoop argue that animacy in general and animacy shifts in particular can be better understood if we take into account the distinction between grammatical and conceptual animacy. They argue that possible mismatches between conceptual and grammatical animacy licence animacy shifts. Their approach seems promising to me,[1] but I would nevertheless like to take this opportunity to discuss further some of the points made in the paper.

If I understand their claims correctly, de Swart and de Hoop propose that linguistic animacy is discrete and binary and conceptual animacy is gradient and non-binary. Linguistic animacy is modelled with binary features and conceptual animacy is modeled with a hierarchy (the animacy hierarchy). I view their proposal as an invitation to consider further the question whether featural analyses are necessarily discrete and analyses that appeal to hierarchies necessarily gradient.

2 Are hierarchies necessarily gradient?

In section 5, de Swart and de Hoop write:

> A view of animacy as discrete types seems directly at odds with the general perception of animacy in the linguistic literature as a gradient notion. Typically, animacy is represented as a category that is ordered along a continuum (or hierarchy)..."

The hierarchy they have in mind is the familiar *human* > *animate* > *inanimate* hierarchy (they cite Comrie 1989 and Aissen 2003 in section 3), where "animate" is understood as non-human animates (that is, animals).

*I am grateful to Raj Singh for providing helpful comments on a previous version of this paper.

[1]In fact, in previous work, Shiva Bayanati and I similarly argue that careful consideration of the tripartite distinction between grammatical animacy, conceptual animacy and biological animacy is crucial for an understanding of verbal agreement marking in Inari Saami and Persian (Bayanati and Toivonen 2015, In review).

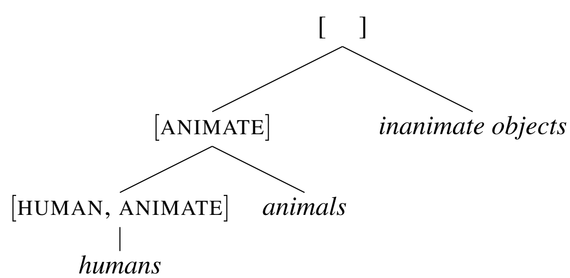
One way of thinking about the hierarchy as implying gradience is to stress that the *animate* category falls between humans and inanimates and is therefore an "in-between" category. Perhaps that can be viewed as a kind of gradience. One can also consider other possible categories, such as higher animals (perhaps pets) and lower animals (mice and maybe insects). Further categories can be added if we include definiteness and the person hierarchy; see Silverstein 1976: 122, for example (but see also Zaenen et al. 2004, Croft 1990 for discussion). In this way, the animacy hierarchy can be thought of as gradient, in a sense, because there are several categories between the highest and the lowest category.

This does not seem to be exactly what de Hoop and de Swart have in mind. When they discuss gradience, they bring up factors such as empathy (citing Yamamoto 1999). The degree of empathy a language user feels towards an entity affects the animacy value assigned to that entity. They also discuss anthropomorphized entities whose animacy value depends on "various physical and psychological properties". This discussion implies that the animacy hierarchy is in fact not neatly divided into human, animate and inanimate. There are no clear borders between the three categories: it is possible for something to fall between human and animate for example.

However, a hierarchy is not necessarily gradient. A hierarchy can be viewed as ordered categories that are discrete. It is possible to formally model something as a hierarchy without allowing for gradience between (or within) the categories on the hierarchy. It would in principle be possible to maintain de Swart and de Hoop's claim that linguistic animacy is discrete and conceptual animacy is gradient, even if the linguistic formalization in some way were to appeal to a hierarchy of categories. One way of doing this would be through harmonic alignment in Optimality Theory, as proposed by Aissen (2003).

It also seems to me that it would be possible to think about animacy in terms of a feature inheritance hierarchy,[2] perhaps along the lines suggested in (1):

(1)



On the highest level, there is no animacy specification. The next level is specified as [ANIMATE], and the lowest level is [HUMAN] and [ANIMATE]. Inanimates are neither [ANIMATE] nor [HUMAN], and are therefore interpreted as inanimate by default (or perhaps unspecified features get a negative value by default). Animals are specified as [ANIMATE] but not [HUMAN] and humans are specified as [ANIMATE] and [HUMAN]. This approach captures the animacy hierarchy intuition that animals hold a place in

[2]For approaches to feature hierarchies in linguistic theory, see, for example, Pollard and Sag (1987, 1994: 36–37) and Harley and Ritter (2002).

between humans and inanimates. It also seems intuititve to model the human category as the most highly specified category and inanimates as unspecified: nouns denoting entities with high animacy are more likely to be morphologically marked for case and agreement, whereas inanimates are more likely to be morphologically unspecified for such features (Aissen 2003). An analysis along these lines would be hierarchical but not gradient.

3 Features and gradience

Although conceptual animacy is gradient, according to de Swart and de Hoop, "linguistically animacy is a discrete phenomenon that can be modelled in terms of binary features [+/- HUMAN] and [+/- ANIMATE]". The authors further take the features to reflect *types*: e_{human}, $e_{animate}$ and $e_{inanimate}$. In this section, I will make the simple point that even if animacy is modelled with features and types in the grammar, linguistic gradience is possible.

Two binary animacy features are proposed by de Swart and de Hoop: [+/- ANIMATE] and [+/- HUMAN]. It is unclear to me whether both animacy features are intended to be available in all languages, or whether a langauge has one feature or the other.

On the one hand, the authors stress that, linguistically, animacy presents itself in a binary way. They claim that this is because linguistic phenomena involving animacy are binary, and they provide the presence or absence of case and pronoun selection as examples. This would imply that a given language has either an [ANIMATE] feature or a [HUMAN] feature: if both features are available there would of course be four possibilities (one of which, [- ANIMATE] and [+ HUMAN], seems impossible). However, there are languages that distinguish between [+/- HUMAN] in the pronominal system and [+/- ANIMATE] in object case marking, so perhaps the binarity restriction is intended to hold for individual phenomena rather than for languages?

On the other hand, de Swart and de Hoop do seem to intend for the two binary features to be considered together. They mention languages where the grammar seems to rely on a tripartite animacy distinction. Of course, in order to capture three categories, more than one binary feature is needed. Also, de Swart and de Hoop's discussion of semantic maps crucially relies on both features: the fact that human and inanimate entities never pattern together to the exclusion of animate (non-human) entities is explained by the fact that human and inanimate entities cannot be described together with a single feature, but humans and animals can ([+ ANIMATE]), and animals and inanimates can ([- HUMAN]). (Of course, the latter point might be intended to be a theoretical point only, not directly pertinent in any one specific language.)

There is a growing literature on grammatical gradience; see, for example, Bresnan and Nikitina (2009) and Aarts (2004) (Aarts 2004 also contains a useful overview of theoretical approaches to gradience). I believe that it would in principle be possible to combine those approaches with a featural analysis of animacy.

Bresnan and Nikitina (2009) analyze the English dative construction with Stochastic Optimality Theory. They model the variation between NP-PP and NP-NP complements of dative verbs with stochastic evaluation of constraints on a scale. If two constraints are close together, the ranking may sometimes switch. A Stochastic OT model could be envisioned for modelling gradience that relates to animacy. For example, there might be a constraint against [- ANIMATE] objects with accusative marking. However, there could be another constraint penalizing the combination of high degree of empathy and an object with oblique case. Consider the following scenario:

- The referent of the object NP is [- ANIMATE] and associated with high empathy.
- The $\boxed{\text{*[- ANIMATE]/Acc}}$ constraint outranks the $\boxed{\text{*high empathy/Obl}}$ constraint.

In this case, the object would receive oblique case marking. However, now consider the following possibility, made available by Stochastic OT if the constraints are close together:

- The referent of the object NP is [- ANIMATE] and associated with high empathy. (Same as above.)
- The $\boxed{\text{*high empathy/Obl}}$ constraint outranks the $\boxed{\text{*[- ANIMATE]/Acc}}$ constraint.

In this case the object receives accusative case marking, even though the referent is inanimate. Now, there might be problems with the particular analysis we have sketched here, but the idea is that it illustrates that gradience in grammar is in principle possible to model even if binary features are adopted to represent animacy.

Aarts (2004) also models linguistic gradience but takes a different approach. He develops a set-theoretic model of gradience which formalizes the concept of prototypes. This allows different entities from the same category to be closer to the prototype of that category than others. Adopting Aarts's theory (or some other characterization of prototypes), it is possible to model gradient animacy in the grammar. We can make use of features to describe human, animate and inanimate entities, but we can still model the fact that some entities are more prototypical members of their groups than others. For example, an inanimate object that we feel empathy for (like a car or a beloved toaster) might be further from the prototype than other inanimates. Similarly, spiders might be further from the [+ANIMATE] prototype than rabbits.

4 Concluding remarks

In these brief comments, I have highlighted some interesting questions that are raised by de Swart and de Hoop's discussion of animacy and animacy shifts. *Shifting Animacy* emphasizes that careful analysis of animacy (along with the observed animacy shifts) can teach us something about the distinction between grammar proper and the conceptualization of the entities that words refer to. Their proposal further invites discussion about the relationship between formal tools and theoretical analyses.

De Swart and de Hoop draw an important distinction between conceptual and grammatical animacy (both of which are different from, but also related to, biological animacy). They argue that conceptual animacy is gradient and naturally modelled by the animacy hierarchy. Grammatical animacy categories, on the other hand, are discrete and modelled by features. I agree that it is natural to model gradience with a hierarchy and discrete categories with features. However, it is in principle possible to model animacy with a hierarchy and still maintain strict boundaries between categories. It is also possible to model animacy with binary features while allowing for gradience and fluidity within and between categories. The formal tools and the theoretical postulates are not intrinsically bound.

References

Aarts, Bas. 2004. Modelling linguistic gradience. *Studies in Language* 28(1): 1–49.

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

Bayanati, Shiva, and Ida Toivonen. 2015. Human, animals, things and animacy features. Talk presented at *Workshop on Animacy in Language and Cognition*, Leeds.

—. In review. Human, animals, things and animacy features. Submitted to Open Linguistics, special issue, 'Effects of animacy in grammar and cognition', D. Nelson and V.-A. Vihman, eds., to appear 2018.

Bresnan, Joan, and Tatiana Nikitina. 2009. The gradience of the dative alternation. In Linda Uyechi and Lian Hee Wee, eds., *Reality Exploration and Discovery: Pattern Interaction in Language and Life*, 161–184. Stanford: CSLI Publications.

Croft, William. 1990. *Typology and Universals*. Cambridge: Cambrige University Press.

Harley, Heidi, and Elizabeth Ritter. 2002. Person and number in pronouns: A feature geometric approach. *Language* 78(3): 482–526.

Pollard, Carl, and Ivan A. Sag. 1987. *Information-Based Syntax and Semantics*. Stanford: CSLI.

—. 1994. *Head-Driven Phrase Structure Grammar*. Stanford, CA: CSLI Publications.

Silverstein, Michael. 1976. Hierarchy of features and ergativity. In Richard Dixon, ed., *Grammatical Categories in Australian Languages*, 112–172. Canberra: Australian Institute of Aboriginal Studies.

Yamamoto, Mutsumi. 1999. *Animacy and reference: a cognitive approach to corpus linguistics*. Amsterdam/Philadelphi: John Benjamins.

Zaenen, Annie, Jean Carletta, Gregory Garretson, Joan Bresnan, Andrew Koontz-Garboden, Tatiana Nikitina, Mary C. O'Connor, and Tom Wasow. 2004. Animacy encoding in English: Why and how. In *Proceedings of the 42nd Annual Meeting of the Association for Computational Linguistics (ACL'04), Workshop on Discourse Annotation*, 118–125. Barcelona.